ARCHIVE

for

RATIONAL MECHANICS

and

ANALYSIS

Edited by

C. TRUESDELL

Volume 5, Number 2

SPRINGER-VERLAG
BERLIN - GÖTTINGEN - HEIDELBERG
(Postverlagsort Berlin · 10. 5. 1960)

Mechanicam vero duplicem Veteres constituerunt: Rationalem *quae per Demonstrationes accurate procedit, & Practicam. Ad practicam spectant Artes omnes Manuales a quibus utique* Mechanica *nomen mutuata est. Cum autem Artifices parum accurate operari soleant, fit ut* Mechanica *omnis a* Geometria *ita distinguatur, ut quicquid accuratum sit ad* Geometriam *referatur, quicquid minus accuratum ad* Mechanicam. *Attamen errores non sunt Artis sed Artificum. Qui minus accurate operatur, imperfectior est Mechanicus, & si quis accuratissime operari posset, hic foret Mechanicus omnium perfectissimus.* NEWTON

La généralité que j'embrasse, au lieu d'éblouïr nos lumieres, nous découvrira plutôt les véritables loix de la Nature dans tout leur éclat, & on y trouvera des raisons encore plus fortes, d'en admirer la beauté & la simplicité. EULER

Ceux qui aiment l'Analyse, verront avec plaisir la Méchanique en devenir une nouvelle branche ... LAGRANGE

The ARCHIVE FOR RATIONAL MECHANICS AND ANALYSIS nourishes the discipline of mechanics as a deductive, mathematical science in the classical tradition and promotes pure analysis, particularly in contexts of application. Its purpose is to give rapid and full publication to researches of exceptional moment, depth, and permanence.

Each memoir must meet a standard of rigor set by the best work in its field Contributions must consist largely in original research; on occasion, an expository paper may be invited.

English, French, German, Italian, and Latin are the languages of the Archive. Authors are urged to write clearly and well, avoiding an excessively condensed or crabbed style.

Manuscripts intended for the Archive should be submitted to an appropriate member of the Editorial Board.

The ARCHIVE FOR RATIONAL MECHANICS AND ANALYSIS appears in numbers struck off as the material reaches the press; five numbers constitute a volume. Subscriptions may be entered through any agent. The price is DM 96.— per volume.

For each paper 75 offprints are provided free of charge.

The Invariance Group for Hamiltonian Systems of Partial Differential Equations

I. Analysis

Dominic Gardiner Bowling Edelen

Communicated by J. L. Ericksen

Contents

Introduction

Many of the disciplines associated with theoretical physics and the allied engineering subjects rely on variational statements which naturally arise from physical considerations or which are judiciously adjusted so as to do so. For those disciplines in which the variational statements lead to ordinary differential equations, the Hamiltonian formalism and the theory of canonical maps provide basic investigatory tools of extreme generality and breadth of application. There are, however, no such basic investigatory tools available when the variational statements lead to partial differential equations. Thus it would be advantageous to investigate the formal structure of partial differential equations which arise from variational statements from the standpoint of extending the Hamiltonian formalism and the theory of canonical maps where existent.

There is, at present, little or no distinction made between the assumptions which delineate the intrinsic geometry of the space of independent variables and the assumptions employed in constructing variational statements concerning collections of functions defined relative to the space of independent variables. This situation is particularly surprising in that one need only assume the space of independent variables is a Hausdorff space in order to obtain a well defined calculus of variations, and hence to construct variational statements. It would thus also be advantageous to investigate the formal structure of partial differential equations which arise from variational statements, where the space of independent variables is assumed to be a Hausdorff space for which there is no intrinsic geometry assumed.

The orientation of this paper is toward an answer to the following question: What are the extensions of the Hamiltonian formalism and the theory of canonical maps to systems of partial differential equations which arise from variational statements, where no assumptions are made concerning the space of independent variables other than that it shall form a Hausdorff space? It is realized that a complete answer to this question is a task of dismayingly large scope. However, it is felt that the principal results are presented from which a complete answer may be developed.

The reader is assumed to be familiar with the general results of Hamiltonian formalism and the theory of canonical maps for systems of ordinary differential equations (henceforth referred to as the classic theory)[1]. Such knowledge is not

[1] For a concise treatment of the results of the classic theory see DE WITT, B. S.: Rev. Mod. Phys. **29**, No. 3, 377–384 (1957). For a more exhaustive treatment see WHITTAKER, E. T.: A Treatise on the Analytical Dynamics of Particles and Rigid Bodies, 4th ed. Cambridge, 1937. WINTNER, A.: The Analytical Foundations of Celestial Mechanics, Chapters I and II. Princeton University Press, 1947.

mandatory in that the results and developments given, when specialized to the case of one independent variable, constitute a direct development of the classic theory insofar as it will be needed for comparison and understanding. In those cases in which the results differ markedly from the results of the classic theory, the differences will be explicitly demonstrated. When it is felt to be illustrative, reference to the corresponding results of the classic theory will be made from the original literature. This organization appears preferable to an introductory exposition of the pertinent literature in that, first, it enables the reader to draw immediate comparisons, and second, there is almost no literature for the case of more than one independent variable in which one of the independent variables is not given preference over the others through functional representations.

Section I. Preliminary Considerations

Notation

An ordered collection of a finite number of scalars b^i will be referred to as the components of a vector $\boldsymbol{b}$. By an n-vector will be meant a vector with n components.

Let $\mathscr{E}_n$ be an n-dimensional vector space of variable n-vectors $\boldsymbol{x}$. This space will be referred to as the space of independent variables.

A point set $\mathscr{D}_n$ contained in $\mathscr{E}_n$ will be referred to as a domain if it is an open, connected, nonvacuous point set. We denote by $\mathscr{D}_n^*$ the closure of $\mathscr{D}_n$. By the boundary of $\mathscr{D}_n^*$ shall be meant the set $\mathscr{D}_n^* \ominus \mathscr{D}_n$, where $\ominus$ stands for the set-theoretic difference.

A scalar or vector function of $\boldsymbol{x}$ is said to be of class C^ν on $\mathscr{D}_n$, where ν is a fixed positive integer, if the function on the domain $\mathscr{D}_n$ under consideration is such that all partial derivatives of order ν or greater exist and are continuous on $\mathscr{D}_n$.

We shall be concerned with collections of functions of both vector and scalar nature defined over $\mathscr{D}_n^*$. Greek indices shall denote elements of the collections of functions considered, and Latin indices shall denote components of vectors defined over $\mathscr{D}_n^*$. Let $\mathrm{q}_\alpha(\boldsymbol{x})$ be a typical collection of scalar functions defined for all $\boldsymbol{x}$ in $\mathscr{D}_n^*$ and for α an element of $\mathscr{A}$, where $\mathscr{A}$ is a finite set of N elements. The set $\mathscr{A}$ is regarded as the index set of such a collection of functions of $\boldsymbol{x}$ and N will be termed the numerosity of the index set $\mathscr{A}$. Such a collection $\mathrm{q}_\alpha(\boldsymbol{x})$ will be denoted by q when there is no possibility of ambiguity. Let q and ${}'\mathrm{q}$ be two collections of functions defined over $\mathscr{D}_n^*$ with the same index set. A map ${}'\mathrm{q} = {}'\mathrm{q}(\mathrm{q}, \boldsymbol{x})$ will be said to be of class $C^{[\nu]}$ if both q and ${}'\mathrm{q}$ are of class C^ν and the map has a local inverse everywhere in the $\mathscr{D}_n$ under consideration. If q and ${}'\mathrm{q}$ are considered as imbedded in spaces of functions of class C^ν, then the map ${}'\mathrm{q} = {}'\mathrm{q}(\mathrm{q}, \boldsymbol{x})$ is said to be of class $C^{[\nu]}$ if it is locally topological.

Partial differentiation will be denoted by a comma. When the partial derivative is to be taken with respect to an element of a collection of functions defined over $\mathscr{D}_n$ or $\mathscr{D}_n^*$, it will be denoted by a comma followed by the symbol for the particular element as a subscript (that is, $H_{,\mathrm{q}} = \partial H / \partial \mathrm{q}$). When differentiation is with respect to a component of $\boldsymbol{x}$, say x^i, it will be denoted by $(\ )_{,i}$. Differentiation of a functions of q and $\boldsymbol{x}$ with respect to x^i where the q

are held constant will be denoted by ∂_i (that is, $\partial_i H = \partial H/\partial x^i|_{\mathrm{q}}$). Hence,

$$H(\mathrm{q}(\boldsymbol{x}), \boldsymbol{x})_{,i} = \sum_{\alpha \in \mathscr{A}} H_{,\mathrm{q}_\alpha} \mathrm{q}_{\alpha,i} + \partial_i H.$$

Whenever a Latin or Greek index appears twice in a term or product, summation is implied. Whenever summation is implied on a Greek index, the summation is to be extended throughout the index set $\mathscr{A}$. The situation in which an index appears twice but is not to be summed will be denoted by $\not\Sigma$ with the index over which summation is not to be extended beneath it (that is, "no summation on α" is written $\underset{\alpha}{\not\Sigma}$).

Column matrices will be denoted by $\{\ \}$, row matrices by $[\ ]$, and square matrices by $\boldsymbol{A}, \boldsymbol{B}, \ldots$. Denote by $\boldsymbol{E}$ the $N \times N$ element identity matrix and by $\boldsymbol{E}^*$ the $N(n+1) \times N(n+1)$ element identity matrix. All matrices considered will be either defined over or relative to the collections of functions defined over $\mathscr{D}_n$ or $\mathscr{D}_n^*$.

Let $\{u\}$ be an S-element column matrix and $\{v(u)\}$ an S-element column matrix each of whose elements is a function of the elements of $\{u\}$. By $\{v\}_{,\{u\}}$ we shall mean the square matrix defined by

$$\{v\}_{,\{u\}} \stackrel{\text{def}}{=} \boldsymbol{W} = (W_{\alpha\beta}) = (\partial v_\alpha/\partial u_\beta).$$

If $f(u)$ is a scalar function whose arguments are the elements of $\{u\}$, then by $f_{,\{u\}}$ we shall mean the row matrix defined by

$$f_{,\{u\}} \stackrel{\text{def}}{=} [A_\alpha] = [\partial f/\partial u_\alpha].$$

Similarly, $\{f_{,u}\}$ is a column matrix defined by

$$\{f_{,u}\} \stackrel{\text{def}}{=} (f_{,\{u\}})'$$

where "prime" denotes the transpose.

Vectrices and Vectrix Algebra

The analyses and proofs involved in the study of the invariance group associated with Hamiltonian systems of partial differential equations becomes extremely cumbersome without the introduction of what we refer to as vectrices and their associated algebra. Loosely speaking, vectrices are ordered collections of matrices, where the ordering is similar to that used in the definition of vectors as ordered collections of scalars. As might be expected from the similarity of ordering between vectrices and vectors, the algebra of vectrices will have many properties in common with the algebra of vectors.

Let X_n be a Euclidean space of n dimensions which is independent of the space $\mathscr{E}_n$ of variable n-vectors $\boldsymbol{x}$. By an **ordering basis** will be meant a fixed collection of constant orthonormal bases vectors $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ which span X_n. The notation $\vec{e}_i$ rather than $\boldsymbol{e}_i$ is used to note specifically the fact that the $\vec{e}_i$ are subsidiary quantities which are independent of any algebraic processes of either vector or scalar nature in $\mathscr{E}_n$.

Definition. Let $\boldsymbol{A}_i$ $(i = 1, \ldots, n)$ be n arbitrary matrices each of which has the same number of rows and columns and whose elements are defined over or relative to $\mathscr{E}_n$. By a **vectrix** $\vec{\boldsymbol{A}}$ will be meant the collection of matrices $\boldsymbol{A}_i$

ordered by an ordering basis $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, as defined by the equation

$$\vec{A} \stackrel{\text{def}}{=} \vec{e}_i \boldsymbol{A}_i.$$

The matrices $\boldsymbol{A}_i$ will be referred to as the **components** of the vectrix $\vec{A}$. It must be clearly borne in mind that the $\vec{e}_i$ are an ordering basis constructed from the subsidiary space X_n and do not operate algebraically with the component matrices $\boldsymbol{A}_i$ (*i.e.*, $\vec{e}_i \boldsymbol{A}_i$ is not to be interpreted in the sense of a product between a vector and a matrix). This condition is easily remembered since the $\vec{e}_i$ are defined from the subsidiary space X_n while all the elements of the component matrices $\boldsymbol{A}_i$ are functions defined over or relative to the space $\mathscr{E}_n$ of variable n-vectors. For instance, $\vec{e}_1 \boldsymbol{E}$ is a vectrix whose first component is the identity matrix, and whose remaining components are $N \times N$ null matrices, $\vec{e}_2 \boldsymbol{E}$ is a vectrix whose second component is the identity matrix and whose remaining components are $N \times N$ null matrices, *etc.*

Definition. Let $\vec{A} = \vec{e}_i \boldsymbol{A}_i$ and $\vec{B} = \vec{e}_i \boldsymbol{B}_i$ be two arbitrary vectrices whose components have the same number of rows and columns. The operations of addition, multiplication by a matrix $\boldsymbol{M}$, transposition, and **inner multiplication** are defined by

$$\vec{A} + \vec{B} \stackrel{\text{def}}{=} \vec{e}_i (\boldsymbol{A}_i + \boldsymbol{B}_i)$$

$$\boldsymbol{M}\vec{A} \stackrel{\text{def}}{=} \vec{e}_i \boldsymbol{M}\boldsymbol{A}_i, \qquad \vec{A}\boldsymbol{M} \stackrel{\text{def}}{=} \vec{e}_i \boldsymbol{A}_i \boldsymbol{M}$$

$$\vec{A}' \stackrel{\text{def}}{=} \vec{e}_i \boldsymbol{A}_i'$$

$$\vec{A} \circ \vec{B} \stackrel{\text{def}}{=} \boldsymbol{A}_i \boldsymbol{B}_i$$

respectively, where $\boldsymbol{M}\boldsymbol{A}_i$ and $\boldsymbol{A}_i \boldsymbol{B}_i$ are multiplied by the laws of matrix multiplication.

Definition. Two vectrices $\vec{A} = \vec{e}_i \boldsymbol{A}_i$ and $\vec{B} = \vec{e}_i \boldsymbol{B}_i$ are said to be equal if and only if the component matrix $\boldsymbol{A}_i$ is equal to the component matrix $\boldsymbol{B}_i$ for all i (*i.e.*, equal component by component).

From the above definitions, it is evident that vectrices admit the following algebraic properties:

$$\vec{A} + \vec{B} = \vec{B} + \vec{A}$$

$$\vec{A} + (\vec{B} + \vec{C}) = (\vec{A} + \vec{B}) + \vec{C}$$

$$\vec{A} + \vec{B} = \vec{A} + \vec{C} \Rightarrow \vec{B} = \vec{C}$$

$$\boldsymbol{M}(\vec{A} + \vec{B}) = \boldsymbol{M}\vec{A} + \boldsymbol{M}\vec{B}$$

$$(\boldsymbol{M} + \boldsymbol{N})\vec{A} = \boldsymbol{M}\vec{A} + \boldsymbol{N}\vec{A}$$

$$(\boldsymbol{M}\boldsymbol{N})\vec{A} = \boldsymbol{M}(\boldsymbol{N}\vec{A})$$

$$\boldsymbol{M}\vec{A} = \vec{A}\boldsymbol{M}$$

if and only if $\boldsymbol{M}$ commutes with $\boldsymbol{A}_i$ for all i

$$(\boldsymbol{M}\vec{A})' = \vec{A}'\boldsymbol{M}'$$

$$\vec{A} \bullet \boldsymbol{M}\vec{B} = \vec{A}\boldsymbol{M} \bullet \vec{B}$$

$$(\vec{A} + \vec{B}) \bullet \vec{C} = \vec{A} \bullet \vec{C} + \vec{B} \bullet \vec{C}$$

$$(\vec{A} \bullet \vec{B})' = \vec{B}' \bullet \vec{A}'.$$

To complete the above results, we define the **null vectrix**, $\vec{\boldsymbol{O}}$, as a vectrix all of whose components are null matrices. We shall refer to vectrices as column, row or square vectrices if their component matrices are column, row or square matrices respectively.

A specific class of vectrices which have central importance in the analyses to follow are those vectrices referred to as structure vectrices.

Definition. A vectrix $\vec{\boldsymbol{A}}$ is said to be a **structure vectrix** if and only if

(i) $\vec{\boldsymbol{A}}$ is a constant square vectrix (1.1)

(ii) $\vec{\boldsymbol{A}}' = -\vec{\boldsymbol{A}}$ (1.2)

(iii) $\det|\vec{\boldsymbol{A}} \bullet \vec{\boldsymbol{A}}| \neq 0.$ (1.3)

Definition. The **right** and **left-hand inverse structure vectrices** corresponding to a structure vectrix $\vec{\boldsymbol{A}}$ are defined by

$$\text{(i)} \qquad (\vec{\boldsymbol{A}}_r^{-1})' = -\vec{\boldsymbol{A}}_l^{-1} \tag{1.4}$$

$$\text{(ii)} \qquad \vec{\boldsymbol{A}}_l^{-1} \bullet \vec{\boldsymbol{A}} = \boldsymbol{E}^* = \vec{\boldsymbol{A}} \bullet \vec{\boldsymbol{A}}_r^{-1} \tag{1.5}$$

$$\text{(iii)} \qquad \vec{\boldsymbol{A}}_r^{-1} = \boldsymbol{C}\vec{\boldsymbol{A}}. \tag{1.6}$$

Theorem 1.1. *The right and left-hand inverse structure vectrices corresponding to a structure vectrix $\vec{\boldsymbol{A}}$ are unique and are given by*

$$\begin{aligned} \vec{\boldsymbol{A}}_l^{-1} &= (\vec{\boldsymbol{A}} \bullet \vec{\boldsymbol{A}})^{-1}\vec{\boldsymbol{A}} \\ \vec{\boldsymbol{A}}_r^{-1} &= \vec{\boldsymbol{A}}(\vec{\boldsymbol{A}} \bullet \vec{\boldsymbol{A}})^{-1}. \end{aligned} \tag{1.7}$$

Proof. Substituting equation (1.6) into (1.5) gives

$$\boldsymbol{E}^* = \vec{\boldsymbol{A}}_l^{-1} \bullet \vec{\boldsymbol{A}} = \boldsymbol{C}\vec{\boldsymbol{A}} \bullet \vec{\boldsymbol{A}},$$

from which we obtain

$$\boldsymbol{C} = (\vec{\boldsymbol{A}} \bullet \vec{\boldsymbol{A}})^{-1},$$

whose existence is assured by equation (1.3) since $\vec{\boldsymbol{A}}$ is a structure vectrix by hypothesis. Thus, by equation (1.6) $\vec{\boldsymbol{A}}_l^{-1} = (\vec{\boldsymbol{A}} \bullet \vec{\boldsymbol{A}})^{-1}\vec{\boldsymbol{A}}$, from which the result $\vec{\boldsymbol{A}}_r^{-1} = \vec{\boldsymbol{A}}(\vec{\boldsymbol{A}} \bullet \vec{\boldsymbol{A}})^{-1}$ follows by equation (1.2) and (1.4). Q.E.D

Given the right and left-hand inverse structure vectrices corresponding to a structure vectrix $\vec{\boldsymbol{A}}$ we may construct collections of vectrices with the properties

$$\text{(i)} \qquad \vec{\boldsymbol{B}} \bullet \vec{\boldsymbol{A}} = \boldsymbol{E}^*$$

$$\text{(ii)} \qquad \vec{\boldsymbol{A}} \bullet \vec{\boldsymbol{F}} = \boldsymbol{E}^*.$$

Specifically, let $\vec{\boldsymbol{C}}_l$ and $\vec{\boldsymbol{C}}_r$ be any two vectrices with the properties $\vec{\boldsymbol{C}}_l \bullet \vec{\boldsymbol{A}} = \boldsymbol{O}$ and $\vec{\boldsymbol{A}} \bullet \vec{\boldsymbol{C}}_r = \boldsymbol{O}$, respectively; then upon setting

$$\vec{\boldsymbol{B}} = \vec{\boldsymbol{A}}_l^{-1} + \vec{\boldsymbol{C}}_l$$

$$\vec{\boldsymbol{F}} = \vec{\boldsymbol{A}}_r^{-1} + \vec{\boldsymbol{C}}_r,$$

we have
$$\vec{B} \bullet \vec{A} = (\vec{A}_l^{-1} + \vec{C}_l) \bullet \vec{A} = \vec{A}_l^{-1} \bullet \vec{A} + \vec{C}_l \bullet \vec{A}$$
$$= \vec{A}_l^{-1} \bullet \vec{A} + O = E^*.$$

The following theorem establishes the basis for many of the results in the sections to follow.

Theorem 1.2. *Let $\vec{A}$ and B be a given structure vectrix and a matrix respectively, and let $\vec{F}$ be a vectrix which is to satisfy*

$$\vec{A} \bullet \vec{F} = B. \tag{1.8}$$

Then the most general form which $\vec{F}$ can assume is

$$\vec{F} = \vec{A}_r^{-1} B + \vec{C}; \quad \vec{A} \bullet \vec{C} = O \tag{1.9}$$

(i.e., $\vec{F}$ is determined by equation (1.8) only to within an arbitrary vectrix whose inner product with $\vec{A}$, from the right, gives the null matrix).

Proof. Multiplying equation (1.8) from the left by $\vec{A}_r^{-1}$ gives

$$\vec{A}_r^{-1}(\vec{A} \bullet \vec{F}) = \vec{A}_r^{-1} B \overset{\text{def}}{=} \vec{G}.$$

Hence

$$\vec{A} \bullet \vec{G} = (\vec{A} \bullet \vec{A}_r^{-1})(\vec{A} \bullet \vec{F}) = \vec{A} \bullet \vec{F},$$

so that

$$\vec{A} \bullet (\vec{F} - \vec{G}) = O,$$

and hence $\vec{F} - \vec{G} = \vec{C}$ where $\vec{A} \bullet \vec{C} = O$. Thus, since $\vec{G} = \vec{A}_r^{-1} B$ we obtain $\vec{F} = \vec{A}_r^{-1} B + \vec{C}$ where $\vec{C}$ is an arbitrary vectrix to within the condition $\vec{A} \bullet \vec{C} = O$. Q.E.D.

We shall often find it convenient to represent a vectrix as a matrix whose elements are linear functions of the ordering base. For example,

$$\begin{pmatrix} \vec{e}_1 & \vec{e}_1 - \vec{e}_2 & 2\vec{e}_3 \\ \vec{e}_2 & \vec{e}_3 & -\vec{e}_1 \\ \vec{e}_3 & \vec{e}_1 & 4\vec{e}_2 \end{pmatrix} = \vec{B}$$

represents the vectrix whose component matrices are given by

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 0 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

respectively; that is

$$\vec{B} = \vec{e}_1 \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} + \vec{e}_2 \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix} + \vec{e}_3 \begin{pmatrix} 0 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Situations will arise in Sections IV and VII in which summations of the form $A_i x^i$ will appear, where A_i are the components of a vectrix and x^i are the components of an n-vector. Since both vectrices and n-vectors are ordered in

a similar manner, we may introduce the following notation:

$$\vec{A} \otimes \boldsymbol{x} \stackrel{\text{def}}{=} \boldsymbol{A}_i \, x^i. \tag{1.10}$$

Since vectrices are ordered relative to an ordering base obtained from the subsidiary space X_n, it is evident that the order of terms in the $\otimes$ product, as defined by equation (1.10), is immaterial; that is

$$\vec{A} \otimes \boldsymbol{x} = \boldsymbol{x} \otimes \vec{A}.$$

Local and Non-Local Questions

To this point, we have made no specific statements concerning the structure of $\mathscr{E}_n$ other than it is a space of ordered collections of scalars, which in no way implies any sort of structure for $\mathscr{E}_n$ other than that for which the algebra of its elements is a linear algebra. We now proceed to delineate the postulates of structure for $\mathscr{E}_n$ which shall be implied in the remainder of this work.

By a space $\mathscr{E}_n$ we shall mean a set of ordered collections of scalars, called points, for which one can define a system of sub-sets called neighborhoods, satisfying the following conditions:

(i) The points of each neighborhood can be put into a one to one reciprocal correspondence with the interior points of a hypersphere of the Euclidean space of n dimensions.

(ii) Each point of $\mathscr{E}_n$ belongs to at least one neighborhood.

(iii) For any two neighborhoods of any given point of $\mathscr{E}_n$ their exists a neighborhood which is contained in the intersection of the two given neighborhoods.

(iv) If a point $\boldsymbol{b}$ is contained in a neighborhood $U(\boldsymbol{a})$, of a point $\boldsymbol{a}$, then there exists a neighborhood $V(\boldsymbol{b})$ which is contained in $U(\boldsymbol{a})$.

(v) If $\boldsymbol{a}$ and $\boldsymbol{b}$ are any two points of $\mathscr{E}_n$, then there exist neighborhoods $U(\boldsymbol{a})$ and $V(\boldsymbol{b})$, of $\boldsymbol{a}$ and $\boldsymbol{b}$ respectively, such that $U(\boldsymbol{a})$ and $V(\boldsymbol{b})$ have no points in common.

Simply stated, we require $\mathscr{E}_n$ to be a Hausdorff[2] space.

In addition to the above structure, the $\mathscr{E}_n$ considered may have a well defined differential geometry for a particular problem under consideration. Whether this is the case or not is a moot question relative to the considerations of this study. We shall require no further knowledge of the $\mathscr{E}_n$ other than its satisfaction of postulates (i) through (v) above.

Let $f_i(\boldsymbol{x})$ $(i = 1, \ldots, n)$ be an ordered collection of functions of class C^1 defined over a given $\mathscr{D}_n^*$ of $\mathscr{E}_n$, and consider the quantity $\mathscr{I}$ defined by

$$\mathscr{I} = \int_{\mathscr{D}_n^*} f_{i,i} \, d v_n \,, \tag{1.11}$$

where $d v_n$ stands for $\prod_{i=1}^{n} d x^i$. Using the ordering basis $\vec{e}_1, \ldots, \vec{e}_n$, we may represent f_i as the components of a vectrix whose component matrices are the 1×1 matrices f_i; that is

$$\vec{f} = \vec{e}_i \, f_i. \tag{1.12}$$

[2] WILDER, L. R.: Topology of Manifolds. Amer. Math. Soc. Colloquium Publ. No. 32 (1949).

Introducing the vectrix operator $\vec{\nabla}$, as defined by

$$\vec{\nabla} \stackrel{\text{def}}{=} \vec{e}_i(\ )_{,i},$$

equation (1.11) takes the form

$$\mathscr{I} = \int_{\mathscr{D}_n^*} \vec{\nabla} \bullet \vec{f}\, d v_n. \tag{1.13}$$

We may, with no loss of generality, view equation (1.13) as resulting from imbedding $\mathscr{D}_n$ in X_n by identifying every point $\boldsymbol{x} = (x^1, \ldots, x^n)$ of $\mathscr{D}_n$ with a point in X_n with the coordinates $(x^1, \ldots, x^n)$, since $\mathscr{E}_n$ and X_n satisfy postulates (i) through (v) and hence are topologically equivalent.

We assume at this point and henceforth that the boundary of $\mathscr{D}_n^*$ is rectifiable. Under this assumption, and noting that both $\mathscr{E}_n$ and X_n are assumed to be Hausdorff spaces, the divergence theorem is applicable. Thus

$$\mathscr{I} = \int_{\mathscr{D}_n^*} \vec{\nabla} \bullet \vec{f}\, d v_n = \int_{\mathscr{D}_n^* \ominus \mathscr{D}_n} \sum_i f_i\, d x^1 \ldots d x^{i-1}\, d x^{i+1} \ldots d x^n. \tag{1.14}$$

Introducing the symbolic vectrix $\vec{N}\, d S$ by the equation

$$\vec{N}\, d S = \vec{e}_i N_i\, d S = \vec{e}_i\, d x^1 \ldots d x^{i-1}\, d x^{i+1} \ldots d x^n, \tag{1.15}$$

equation (1.14) takes the symbolic form

$$\mathscr{I} = \int_{\mathscr{D}_n^*} \vec{\nabla} \bullet \vec{f}\, d v_n = \int_{\mathscr{D}_n^* \ominus \mathscr{D}_n} \vec{f} \bullet \vec{N}\, d S. \tag{1.16}$$

Using only the topological structure of $\mathscr{E}_n$ and X_n and the ordering base, we have obtained the global result stated in equation (1.16). Since equation (1.16) is the basis for obtaining conservation laws, we shall be able to obtain conservation laws for the systems of equations studied without requiring any knowledge of a differential geometry for the space $\mathscr{E}_n$. Although this result is an obvious one in that the divergence theorem is a topological theorem, it will be of considerable moment in the second of this series of papers. For this reason the above detailed discussion is included.

Section II. Canonical Maps and their Fundamental Characterization

Hamiltonian Systems and Canonical Maps

Let $q_\alpha(\boldsymbol{x})$ be a collection of scalar functions of class C^2 and $p_{\alpha i}(\boldsymbol{x})$ be a collection of n-vector functions of class C^1 for all $\boldsymbol{x}$ in some given $\mathscr{D}_n^*$, where α ranges over a finite index set $\mathscr{A}$ of numerosity N. Denote such a collection of functions by (p, q). Let $\mathfrak{H}$ be the space of all functions of class C^2 in their $N + nN + n$ arguments q_α, $p_{\alpha i}$, $\boldsymbol{x}$. The space $\mathfrak{H}$ will be referred to as the **space of admissible bases functions**. A collection of functions (p, q) is said to form a **Hamiltonian system base** $\boldsymbol{H}$ in $\mathscr{D}_n^*$ if (p, q) satisfy the system of partial differential equations

$$\left.\begin{aligned} p_{\alpha i, i} &= -H_{, q_\alpha} \\ q_{\alpha, i} &= H_{, p_{\alpha i}} \end{aligned}\right\} H \in \mathfrak{H} \tag{2.1}$$

for all $\boldsymbol{x}$ in $\mathscr{D}_n^*$. The above equations are the partial differential form of the celebrated Hamiltonian equations which were introduced by HAMILTON[3] in 1834, as will be shown in Section III.

Let $\mathscr{T}$ be the collection of all topological maps of class $C^{[1]}$ in (p, q) and of class C^2 in $\boldsymbol{x}$ for all $\boldsymbol{x}$ in $\mathscr{D}_n^*$

$$\mathscr{T}: (p_{\alpha i} = p_{\alpha i}(P_{\beta j}, Q_\beta, \boldsymbol{x});\ \mathrm{q}_\alpha = \mathrm{q}_\alpha(P_{\beta j}, Q_\beta, \boldsymbol{x})). \tag{2.2}$$

If equations (2.1) be written in terms of the new variables (P, Q) for an arbitrary element of the collection $\mathscr{T}$, the form of equations (2.1) will not, in general, be preserved. The concern of this work will be the determination of the subcollection of $\mathscr{T}$ for which the form of equations (2.1) is preserved and the properties of this subcollection.

Definition. A map T of the collection $\mathscr{T}$ will be said to be **canonical** if and only if there exists, for every $H(p, \mathrm{q}, \boldsymbol{x}) \in \mathfrak{H}$, a $\mathrm{K}(P, Q, \boldsymbol{x}) \in \mathfrak{H}$ such that *if* (p, q) form a Hamiltonian system base H in $\mathscr{D}_n^*$, then (P, Q) form a Hamiltonian system base K in $\mathscr{D}_n^*$.

From this definition we have the following statement: The form of equations (2.1) is invariant under the collection of all canonical maps. It should be noted that there exist maps of the collection $\mathscr{T}$ for which the form of equations (2.1) is invariant for a particular $H \in \mathfrak{H}$ *but not for all* $H \in \mathfrak{H}$.

Examining the definition of a canonical map in detail, we see that T is a canonical map if and only if, for every $H \in \mathfrak{H}$, there exists a $\mathrm{K} \in \mathfrak{H}$ with the property: *if* (p, q) form a Hamiltonian system base H, then (P, Q) form a Hamiltonian system base K. Let $\mathfrak{H}_0$ be the subcollection of $\mathfrak{H}$ whose elements, H_0, are such that there exist no (p, q) which form Hamiltonian systems base H_0 (that is, there exist no solutions for (p, q) to equations (2.1) for H_0 and element of $\mathfrak{H}_0$). Denote by $\mathscr{T}_0$ the collection of all elements T_0 of $\mathscr{T}$ which are such that for every element of $\mathfrak{H}_0$, there exists a $\mathrm{K} \in \mathfrak{H}$ with the property: *if* (p, q) form a Hamiltonian system base $H_0 \in \mathfrak{H}_0$, then (P, Q) form a Hamiltonian system base K. Since there are no (p, q) which form Hamiltonian systems base $H_0 \in \mathfrak{H}_0$ by the definition of $\mathfrak{H}_0$, the conditional statement in the above definition of $\mathscr{T}_0$ is vacuous, and hence we may pick any $\mathrm{K} \in \mathfrak{H}$ and the condition will be satisfied. Hence there are no restrictions on $\mathscr{T}_0$, so that $\mathscr{T}_0$ is identical with $\mathscr{T}$. Now, let $\mathscr{T}_1$ be the subcollection of $\mathscr{T}$ whose elements are such that for all elements of $\mathfrak{H} \ominus \mathfrak{H}_0$ there exists a $\mathrm{K} \in \mathfrak{H}$ with the property: *if* (p, q) form a Hamiltonian system base $H \in \mathfrak{H} \ominus \mathfrak{H}_0$, then (P, Q) form a Hamiltonian system base K, and let $\mathscr{T}_c$ be the collection of all canonical maps. From the definition of a canonical map, we have $\mathscr{T}_c = \mathscr{T}_0 \cap \mathscr{T}_1$ since $\mathfrak{H} = (\mathfrak{H} \ominus \mathfrak{H}_0) \cup \mathfrak{H}_0$. We have shown, however, that $\mathscr{T}_0$ is identical with $\mathscr{T}$ so that $\mathscr{T}_1$ is contained in $\mathscr{T}_0$ from whence $\mathscr{T}_0 \cap \mathscr{T}_1 = \mathscr{T}_1$ and hence $\mathscr{T}_c = \mathscr{T}_1$. It is also evident, for all elements of $\mathscr{T}_1$, that if $H \in \mathfrak{H} \ominus \mathfrak{H}_0$, then K must be an element of $\mathfrak{H} \ominus \mathfrak{H}_0$ since the conditional statement in the definition of $\mathscr{T}_1$ requires (P, Q) to form a Hamiltonian system base K. Thus we have the following equivalent reduced form of the definition of a canonical map.

Definition. A map T of the collection $\mathscr{T}$ will be said to be **canonical** if and only if there exists, for every $H \in \mathfrak{H} \ominus \mathfrak{H}_0$, a $\mathrm{K} \in \mathfrak{H} \ominus \mathfrak{H}_0$ such that if (p, q) form

[3] HAMILTON, W. R.: Phil. Transactions 95 (1835).

a Hamiltonian system base H in $\mathscr{D}_n^*$, then (P, Q) form a Hamiltonian system base K in $\mathscr{D}_n^*$.

Before using the above definition to characterize the collection of canonical maps analytically, we proceed to introduce matrix notations which will be of great facility in the application of the above definition and in subsequent work. Define the vectrices $\vec{\boldsymbol{E}}_2$ and $\vec{\boldsymbol{E}}_3$ by

$$\vec{\boldsymbol{E}}_2 = [\vec{e}_1 \boldsymbol{E}\ \vec{e}_2 \boldsymbol{E} \ldots \vec{e}_n \boldsymbol{E}]; \qquad \vec{\boldsymbol{E}}_3 = -\vec{\boldsymbol{E}}_2',$$

where $\boldsymbol{E}$ is the $N \times N$ identity matrix and the prime denotes transposition. Let $\vec{\boldsymbol{O}}_1$ and $\vec{\boldsymbol{O}}_2$ be $N \times N$ and $nN \times nN$ null vectrices respectively and define the constant vectrix $\vec{\boldsymbol{I}}$ by

$$\vec{\boldsymbol{I}} = \begin{pmatrix} \vec{\boldsymbol{O}}_1 & \vec{\boldsymbol{E}}_2 \\ \vec{\boldsymbol{E}}_3 & \vec{\boldsymbol{O}}_2 \end{pmatrix} \tag{2.3}$$

so that

$$\vec{\boldsymbol{I}} = \left(\begin{array}{c|cccc} \vec{\boldsymbol{O}}_1 & \vec{e}_1 \boldsymbol{E} & \vec{e}_2 \boldsymbol{E} & \ldots & \vec{e}_n \boldsymbol{E} \\ \hline -\vec{e}_1 \boldsymbol{E} & & & & \\ -\vec{e}_2 \boldsymbol{E} & & \vec{\boldsymbol{O}}_2 & & \\ \vdots & & & & \\ -\vec{e}_n \boldsymbol{E} & & & & \end{array}\right).$$

From the definition of $\vec{\boldsymbol{I}}$ it is evident that $\vec{\boldsymbol{I}}$ satisfies the conditions required of a structure vectrix; that is,

(i) $\vec{\boldsymbol{I}}$ is a constant square vectrix

(ii) $\vec{\boldsymbol{I}}' = -\vec{\boldsymbol{I}}$ (2.4)

(iii) $\det |\vec{\boldsymbol{I}} \bullet \vec{\boldsymbol{I}}| = (-1)^{N(n+1)}\, n^N \neq 0.$

By Theorem 1.1 we may construct the right and left-hand inverse structure vectrices corresponding to $\vec{\boldsymbol{I}}$ which have the property

$$\vec{\boldsymbol{I}}_l^{-1} \bullet \vec{\boldsymbol{I}} = \boldsymbol{E}^*; \qquad \vec{\boldsymbol{I}} \bullet \vec{\boldsymbol{I}}_r^{-1} = \boldsymbol{E}^*. \tag{2.5}$$

By the same theorem, we have

$$\vec{\boldsymbol{I}}_l^{-1} = -\vec{\boldsymbol{I}}_r^{-1\prime} = -(\eta)\,\vec{\boldsymbol{I}} \tag{2.6}$$

where by equation (1.6)

$$-(\eta) = (\vec{\boldsymbol{I}} \bullet \vec{\boldsymbol{I}})^{-1} = \boldsymbol{C}.$$

Expanding, we have

$$(\eta) = \left(\begin{array}{c|cccc} \frac{1}{n}\boldsymbol{E} & \boldsymbol{O} & \ldots & & \boldsymbol{O} \\ \hline \boldsymbol{O} & \boldsymbol{E} & \boldsymbol{O} & & \\ \vdots & \boldsymbol{O} & \boldsymbol{E} & \ddots & \\ & & & \ddots & \boldsymbol{O} \\ \boldsymbol{O} & & & \boldsymbol{O} & \boldsymbol{E} \end{array}\right), \tag{2.7}$$

so that

$$\vec{\boldsymbol{I}}_r^{-1} = \left(\begin{array}{c|ccc} \vec{\boldsymbol{O}}_1 & -\vec{e}_1 \boldsymbol{E} & \dots & -\vec{e}_n \boldsymbol{E} \\ \hline \frac{\vec{e}_1 \boldsymbol{E}}{n} & & & \\ \vdots & & \vec{\boldsymbol{O}}_2 & \\ \frac{\vec{e}_n \boldsymbol{E}}{n} & & & \end{array}\right). \tag{2.8}$$

Care must be exercised in the use of the right and left inverse vectrices. For instance if $\vec{\boldsymbol{A}}$ and $\vec{\boldsymbol{B}}$ are vectrices which are related by the equation

$$\vec{\boldsymbol{I}}(\vec{\boldsymbol{I}}_l^{-1} \bullet \vec{\boldsymbol{A}}) = \vec{\boldsymbol{B}}$$

for given $\vec{\boldsymbol{A}}$, then

$$\vec{\boldsymbol{I}}_l^{-1} \bullet \vec{\boldsymbol{I}}(\vec{\boldsymbol{I}}_l^{-1} \bullet \vec{\boldsymbol{A}}) = \vec{\boldsymbol{I}}_l^{-1} \bullet \vec{\boldsymbol{B}}$$

gives

$$\vec{\boldsymbol{I}}_l^{-1} \bullet \vec{\boldsymbol{A}} = \vec{\boldsymbol{I}}_l^{-1} \bullet \vec{\boldsymbol{B}}$$

which would appear to imply that $\vec{\boldsymbol{A}} = \vec{\boldsymbol{B}}$. This is not the case in general as shown in Theorem 1.2. A specific result which will be needed is as follows: if $\{\vec{\mathrm{K}}\}$ and $\{\vec{L}\}$ are column vectrices where $\{\vec{\mathrm{K}}\}$ is given, then $\vec{\boldsymbol{I}}(\vec{\boldsymbol{I}}_l^{-1} \bullet \{\vec{\mathrm{K}}\}) = \{\vec{L}\}$ implies $\{\vec{L}\} = \{\vec{\mathrm{K}}\}$ only if

$$\{\vec{\mathrm{K}}\} = \begin{Bmatrix} a_{\alpha i} \vec{e}_i \\ a_{\alpha 1 i} \vec{e}_i \\ \vdots \\ a_{\alpha n i} e_i \end{Bmatrix}, \qquad a_{\alpha i j} = b_\alpha \delta_{ij}. \tag{2.9}$$

Let $\{R\}$ be the $N(n+1)$ element column matrix defined by

$$\{R\} = \begin{Bmatrix} \{q_\alpha\} \\ \{p_{\alpha 1}\} \\ \vdots \\ \{p_{\alpha n}\} \end{Bmatrix} \tag{2.10}$$

and let $\vec{\nabla}$ denote the operator defined by

$$\vec{\nabla} \overset{\text{def}}{=} \vec{e}_i \frac{\partial}{\partial x_i}, \tag{2.11}$$

so that

$$\vec{\nabla}\{R\} = \vec{e}_i \begin{Bmatrix} \{q_{\alpha, i}\} \\ \{p_{\alpha 1, i}\} \\ \vdots \\ \{p_{\alpha n, i}\} \end{Bmatrix}$$

is a well defined vectrix.

With the above matrix representations, the Hamiltonian equations (equations (2.1)) are given by

$$\vec{\boldsymbol{I}} \bullet \vec{\nabla}\{R\} + \{H_{,R}\} = \{0\} \tag{2.12}$$

(that is,

$$\left\{\begin{matrix} \{p_{\alpha i,i}\} \\ \{-q_{\alpha,1}\} \\ \vdots \\ \{-q_{\alpha,n}\} \end{matrix}\right\} + \left\{\begin{matrix} \{H_{,q\alpha}\} \\ \{H_{,p\alpha 1}\} \\ \vdots \\ \{H_{,p\alpha n}\} \end{matrix}\right\} = \{0\}\,)$$

as seen by comparison with equations (2.1).

It should be noted that vectrices may be equally well used to represent tensorial equations or non-tensorial equations. To illustrate the method we consider the following cases.

Case 1. Let I^i and φ be a contravariant tensor of of rank one and a scalar density, respectively, which satisfy the equations

$$I^i_{;i} = 0; \qquad \varphi_{;i} = 0$$

where (;) is used to represent the covariant derivative. Expanding the indicated covariant derivatives, assuming that the underlying metric space is a Riemannian space with affine connection Γ^i_{jk}, we have

$$I^i_{,i} + \Gamma^i_{ij} I^j = 0, \qquad \varphi_{,k} - \Gamma^i_{ik}\,\varphi = 0$$

which can be represented as follows:

$$\{R\} = \left\{\begin{matrix} \varphi \\ I^1 \\ \vdots \\ I^n \end{matrix}\right\} \qquad \begin{matrix} \vec{\boldsymbol{I}} \circ \vec{\nabla}\{R\} + \{H_{,R}\} = \{0\}, \\ H = \Gamma^i_{ij} I^j \varphi. \end{matrix}$$

Case 2. Let I^i and φ be a contravariant tensor of rank one and a scalar, respectively, which satisfy the system of equations

$$I^i_{;i} = \varphi, \qquad \varphi_{;k} = g_{ki} I^i;$$

then we may represent them by

$$\{R\} = \left\{\begin{matrix} \varphi \\ I^1 \\ \vdots \\ I^n \end{matrix}\right\} \qquad \begin{matrix} \vec{\boldsymbol{I}} \bullet \vec{\nabla}^*\{R\} + \{H_{,R}\} = \{0\} \\ H = \tfrac{1}{2} g_{ij} I^i I^j - \tfrac{1}{2}\varphi^2 \\ \vec{\nabla}^* = \vec{e}_i\,(\ )_{;i}\,. \end{matrix}$$

Alpha Systems

The systems of partial differential equations thus far considered (Hamiltonian Systems) may be generalized to a larger class as follows: A column matrix $\{R\}$ is said to form an **alpha system** with structure vectrix $\vec{\boldsymbol{\alpha}}$ in a given $\mathscr{D}_n^*$ if $\{R\}$ satisfies, in $\mathscr{D}_n^*$, the equation

$$\vec{\nabla}\{R\}' \bullet \vec{\boldsymbol{\alpha}} + H_{,\{R\}} = [0] \tag{2.13a}$$

or

$$\vec{\boldsymbol{\alpha}}' \bullet \vec{\nabla}\{R\} + \{H_{,R}\} = \{0\}. \tag{2.13b}$$

Many of the theorems to follow use only the properties of $\vec{I}$ which are the same as the properties of structure vectrices and do not depend on the particular form of $\vec{I}$. Thus, if we replace the term "Hamiltonian system" in such theorems by the term "alpha system" and replace $\boldsymbol{I}$ and its right and left-hand inverse vectrices by the structure vectrix $\boldsymbol{\alpha}$ and its associated right and left-hand inverse structure vectrices, the theorems will carry over directly. We shall mark all theorems or parts of theorems which carry over directly to alpha systems by asterisks.

As will be seen, the structure vectrix of an alpha system effectively characterizes a majority of the formal properties of such a system. Thus, the reduction of systems of partial differential equations to alpha systems and thereby defining the structure vectrices for such systems gives a convenient method of catagorizing such systems as to their formal properties. In this sense equations (2.13 a) or (2.13 b) may be thought of as defining a canonical form.

Statement of the Problem

Consider a map T of the collection $\mathscr{T}$ $\big(i.e.: T:(p, \mathrm{q}) \to (P, Q)\big)$ which may be represented as

$$T:\{R\} \to \{\tilde{R}\}.$$

By the **associated Jacobian matrix** of T will be meant the matrix

$$\boldsymbol{M} = \{R\}_{,\{\tilde{R}\}} \tag{2.14}$$

which may be written in the alternate form

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} \tag{2.15}$$

where $\boldsymbol{A}$, defined by

$$\boldsymbol{A} = (\mathrm{q}_{\alpha, Q_\beta}), \tag{2.16}$$

is an $N \times N$ matrix; $\boldsymbol{B}$, defined by

$$\boldsymbol{B} = [(\mathrm{q}_{\alpha, P_{\beta 1}}) \cdots (\mathrm{q}_{\alpha, P_{\beta n}})], \tag{2.17}$$

is an $N \times nN$ matrix; $\boldsymbol{C}$, defined by

$$\boldsymbol{C}' = [(p_{\alpha 1, Q_\beta})' \cdots (p_{\alpha n, Q_\beta})'], \tag{2.18}$$

is an $nN \times N$ matrix; and $\boldsymbol{D}$, defined by

$$\boldsymbol{D} = \begin{pmatrix} (p_{\alpha 1, P_{\beta 1}}) & \cdots & (p_{\alpha 1, P_{\beta n}}) \\ \vdots & & \vdots \\ (p_{\alpha n, P_{\beta 1}}) & \cdots & (p_{\alpha n, P_{\beta n}}) \end{pmatrix}, \tag{2.19}$$

is an $nN \times nN$ matrix.

We are now in a position to develop the analytic characterization of canonical maps. Consider the forms which the terms in equation (2.12) take as the result of a map of the collection $\mathscr{T}$. Examining the second term of equation (2.12), we have, for an arbitrary element of $\mathscr{T}$, upon taking the transpose

$$H_{,\{R\}} = H_{,\{\tilde{R}\}} \{\tilde{R}\}_{,\{R\}};$$

but

$$\{\widetilde{R}\}_{,\{R\}} = \boldsymbol{M}^{-1}$$

and

$$\{H_{,R}\} = (H_{,\{R\}})'$$

so that

$$\{H_{,R}\} = \boldsymbol{M}^{-1\prime}\{H_{,\widetilde{R}}\}.$$

Let $\underset{0}{\vec{\nabla}}$ be the vectrix operator defined by

$$\underset{0}{\vec{\nabla}} \overset{\text{def}}{=} \vec{e}_i \partial_i(\ ). \tag{2.20}$$

The quantity $\underset{0}{\vec{\nabla}}\{\widetilde{R}\}$ will be interpreted as the $N(n+1)$ element column vectrix obtained if one evaluates $\underset{0}{\vec{\nabla}}\{\widetilde{R}(R, \boldsymbol{x})\}$ and then expresses R in terms of $\widetilde{R}$ and $\boldsymbol{x}$ by the transformation equations. Considering the first term of equation (2.12), we have under T

$$\vec{\nabla}\{\widetilde{R}\} = \{\widetilde{R}\}_{,\{R\}}\vec{\nabla}\{R\} + \underset{0}{\vec{\nabla}}\{\widetilde{R}\}$$

so that

$$\vec{\nabla}\{R\} = \boldsymbol{M}\,\vec{\nabla}\{\widetilde{R}\} - \boldsymbol{M}\,\underset{0}{\vec{\nabla}}\{\widetilde{R}\}.$$

Hence, the left member of equation (2.12) becomes under T

$$\vec{\boldsymbol{I}}\bullet\vec{\nabla}\{R\} + \{H_{,R}\} = \boldsymbol{M}^{-1\prime}\{\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M}\bullet(\vec{\nabla}\{\widetilde{R}\} - \underset{0}{\vec{\nabla}}\{\widetilde{R}\}) + \{H_{,\widetilde{R}}\}\}, \tag{2.21}$$

upon noting that

$$\vec{\boldsymbol{I}}\bullet\boldsymbol{M}\,\vec{\nabla} = \vec{\boldsymbol{I}}\boldsymbol{M}\bullet\vec{\nabla}.$$

Since the left-hand side of the above equation vanishes by equation (2.12), if (p, q) form a Hamiltonian system and $\boldsymbol{M}^{-1\prime}$ is nonsingular for all $\boldsymbol{x}$ in $\mathscr{D}_n^*$, we have

$$\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M}\bullet(\vec{\nabla}\{\widetilde{R}\} - \underset{0}{\vec{\nabla}}\{\widetilde{R}\}) + \{H_{,\widetilde{R}}\} = \{0\}. \tag{2.22}$$

By definition, T, which results in equation (2.22), is a canonical map if and only if there exists, for every $H \in \mathfrak{H} \ominus \mathfrak{H}_0$, a $\mathrm{K} \in \mathfrak{H} \ominus \mathfrak{H}_0$ such that (P, Q) form a Hamiltonian system base K in $\mathscr{D}_n^*$. Stating that (P, Q) form a Hamiltonian system base K in $\mathscr{D}_n^*$ is equivalent to stating that (P, Q) satisfy

$$\vec{\boldsymbol{I}}\bullet\vec{\nabla}\{\widetilde{R}\} + \{\mathrm{K}_{,\widetilde{R}}\} = \{0\} \tag{2.23}$$

by equation (2.12). We are thus faced with the problem of establishing what conditions must be placed on a map T of the collection $\mathscr{T}$ such that equation (2.21) or (2.22) reduces to equation (2.23) whenever (p, q) form a Hamiltonian system base H for all $H \in \mathfrak{H} - \mathfrak{H}_0$.

The Fundamental Theorem

The arguments leading to the establishment of the required conditions in order for a map to be canonical are somewhat lengthy and involved. For this reason we give, at this point, a cursory outline of the procedure so that the forest may not be lost for the trees.

We first reformulate the problem in terms of operators over matrix collections of function spaces and interpret these operators as inducing mappings to other matrix collections of function spaces. We then fix the basis function and formulate in Lemma 2.1 the conditions that equation (2.21) reduces to equation (2.23) in terms of the kernels of the maps defined over the function spaces considered. We then obtain in Lemma 2.2 necessary and sufficient conditions that the conditions of Lemma 2.1 be satisfied in terms of the existence of subsidiary maps on the operators considered. These two lemmas then allow us to state the necessary and sufficient conditions that equation (2.21) is equivalent to equation (2.23) for fixed H, Lemma 2.3 and Theorem 2.2. The result thus obtained is then required to hold for all basis functions of the collection $\mathfrak{H} \ominus \mathfrak{H}_0$ from which we obtain the fundamental theorem (Theorem 2.2) which analytically characterizes the most general canonical map.

Denote by $\mathfrak{R}$ the space of all $N(n+1)$ element column matrices $\{R(\boldsymbol{x})\}$ of class C^1 for all $\boldsymbol{x}$ in some given $\mathscr{D}_n^*$ and by $\mathfrak{S}$ the space of all $N(n+1)$ element column matrices $\{s(\boldsymbol{x})\}$ of class C for all $\boldsymbol{x}$ in $\mathscr{D}_n^*$. Under the assumed continuity of the elements of $\mathfrak{R}$ we may construct a continuous matrix operator $\{\mathfrak{O}_H\}$ over $\mathfrak{R}$, defined by

$$\begin{gathered}\{\mathfrak{O}_H\}(r) \overset{\text{def}}{=} \vec{\boldsymbol{I}} \bullet \vec{\nabla}\{r\} + \{H_{,r}\} \\ \{r\} \in \mathfrak{R}, \quad H \in \mathfrak{H} \ominus \mathfrak{H}_0, \quad \boldsymbol{x} \in \mathscr{D}_n^*\end{gathered} \tag{2.24}$$

which assigns to every element of $\mathfrak{R}$ a new element which is at least of class C. We consider this operator as a continuous mapping of $\mathfrak{R}$ into $\mathfrak{S}$:

$$\{\mathfrak{O}_H\} : \mathfrak{R} \overset{\text{onto}}{\longrightarrow} \mathfrak{S}_H < \mathfrak{S}.$$

Denote by ϱ_H the kernel of the above map which is assumed to be nonvoid (*i.e.*, $H \in \mathfrak{H} \ominus \mathfrak{H}_0$)

$$\{\mathfrak{O}_H\} : \varrho_H \to \{\Phi\} \tag{2.25}$$

where $\{\Phi\}$ is the zero element of $\mathfrak{S}$ defined by $\begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$. Let T be a map of the collection $\mathscr{T}$ so that

$$T : \mathfrak{R} \to \tilde{\mathfrak{R}}. \tag{2.26}$$

Under such a map we have seen that the operator $\{\mathfrak{O}_H\}$ is transformed by the equation

$$T : \{\mathfrak{O}_H\} \to \{\tilde{\mathfrak{O}}_H\}$$

where

$$\{\tilde{\mathfrak{O}}_H\}(r) \overset{\text{def}}{=} \boldsymbol{M}^{-1\prime}\{\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} \bullet (\vec{\nabla}\{r\} - \underset{0}{\vec{\nabla}}\{r\}) + \{H_{,r}\}\}; \quad \{r\} \in \tilde{\mathfrak{R}} \tag{2.27}$$

$$\left(i.e., \quad \vec{\boldsymbol{I}} \bullet \vec{\nabla}\{R\} + \{H_{,R}\} = \boldsymbol{M}^{-1\prime}\{\boldsymbol{M}'\boldsymbol{I}\boldsymbol{M} \bullet (\vec{\nabla}\{\tilde{R}\} - \underset{0}{\vec{\nabla}}\{\tilde{R}\}) + \{H_{,\tilde{R}}\}\}\right).$$

Since T is the element of the collection $\mathscr{T}$, the elements of $\tilde{\mathfrak{R}}$ are of class C^1, and hence $\{\tilde{\mathfrak{O}}_H\}$ may be interpreted as mapping $\tilde{\mathfrak{R}}$ into $\mathfrak{S}$:

$$\{\tilde{\mathfrak{O}}_H\} : \tilde{\mathfrak{R}} \overset{\text{onto}}{\longrightarrow} \tilde{\mathfrak{S}}_H < \mathfrak{S}.$$

Let $\tilde{\varrho}_H$ be the image of ϱ_H under T, so that

$$\{\tilde{\mathfrak{O}}_H\} : \tilde{\varrho}_H \to \{\Phi\}.$$

We define a new operator $\{\mathfrak{O}_H^*\}$ by

$$\{\mathfrak{O}_H^*\}(u) \underset{\tilde{\mathfrak{R}}}{\overset{\text{def}}{=}} \boldsymbol{M}'\{\mathfrak{O}_H\}(u) \tag{2.28a}$$

(*i.e.*,

$$\{\mathfrak{O}_H^*\}(u) = \boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} \circ (\vec{V}\{u\} - \underset{0}{\vec{V}}\{u\}) + \{H_{,u}\} \tag{2.28b}$$

as is seen from equation (2.27) upon noting that $\boldsymbol{M}$ is nonsingular by hypothesis). The operators $\{\tilde{\mathfrak{O}}_H\}$ and $\{\mathfrak{O}_H^*\}$ both have the same kernel, namely $\tilde{\varrho}_H$, when considered as mapping $\tilde{\mathfrak{R}}$ into $\mathfrak{S}$, as is immediate from equations (2.27) and (2.28) upon noting again that $\boldsymbol{M}$ is nonsingular.

Consider now an operator $\{\mathfrak{O}_{\mathrm{K}}\}$ defined over $\tilde{\mathfrak{R}}$ by

$$\{\mathfrak{O}_{\mathrm{K}}\}(u) \overset{\text{def}}{=} \vec{\boldsymbol{I}} \circ \vec{V}\{u\} + \{\mathrm{K}_{,u}\}, \qquad \{u\} \in \tilde{\mathfrak{R}}, \quad \mathrm{K} \in \mathfrak{H} \ominus \mathfrak{H}_0. \tag{2.29}$$

Since K is assumed to be an element of $\mathfrak{H} \ominus \mathfrak{H}_0$ and the elements of $\tilde{\mathfrak{R}}$ are of class C^1, $\{\mathfrak{O}_{\mathrm{K}}\}$ may be considered as a continuous map of $\tilde{\mathfrak{R}}$ into $\mathfrak{S}$:

$$\{\mathfrak{O}_{\mathrm{K}}\} : \tilde{\mathfrak{R}} \xrightarrow{\text{onto}} \mathfrak{S}_{\mathrm{K}} \subset \mathfrak{S}. \tag{2.30}$$

Denote by ϱ_{K} the kernel of $\{\mathfrak{O}_{\mathrm{K}}\}$ so that

$$\{\mathfrak{O}_{\mathrm{K}}\} : \varrho_{\mathrm{K}} \to \{\Phi\}. \tag{2.31}$$

From the form of equation (2.29) we see that the elements of ϱ_{K} form Hamiltonian system base K in $\tilde{\mathfrak{R}}$. Similarly, from the definition of $\tilde{\varrho}_H$ (that is, $\tilde{\varrho}_H$ is the image of the kernel of $\{\mathfrak{O}_H\}$, in $\tilde{\mathfrak{R}}$, under T) and equation (2.24) it is evident that the elements of $\tilde{\varrho}_H$ are the images, under T, of those elements of $\mathfrak{R}$ which form Hamiltonian systems base H.

Let H be a fixed element of $\mathfrak{H} \ominus \mathfrak{H}_0$. If, for some $\mathrm{K} \in \mathfrak{H} \ominus \mathfrak{H}_0$, $\tilde{\varrho}_H$ is contained in ϱ_{K}, then T is such that it maps every point of $\mathfrak{R}$ which forms a Hamiltonian system base H into a point in $\tilde{\mathfrak{R}}$ which forms a Hamiltonian system base K (that is, $\tilde{\varrho}_H$ contained in ϱ_{K} implies that for every $\{u\}$ an element of $\tilde{\varrho}_H$, $\{\mathfrak{O}_{\mathrm{K}}\}(u) = \{\Phi\}$, and hence $\{u\}$ forms a Hamiltonian system base K). Conversely, if T is to be such that it maps every point of $\mathfrak{R}$ which forms a Hamiltonian system base H into a point in $\tilde{\mathfrak{R}}$ which forms a Hamiltonian system base K, then K must be such that $\{\mathfrak{O}_{\mathrm{K}}\}$ maps $\tilde{\varrho}_H$ into $\{\Phi\}$ and hence $\tilde{\varrho}_H$ must be contained in ϱ_{K}. We have thus proved the following

***Lemma 2.1.** *A necessary and sufficient condition that a map T of the collection $\mathscr{T}$ map all points of $\mathfrak{R}$ which form Hamiltonian systems base H, for fixed $H \in \mathfrak{H} \ominus \mathfrak{H}_0$, into Hamiltonian systems in $\tilde{\mathfrak{R}}$ is that there exist at least one function $\mathrm{K} \in \mathfrak{H} \ominus \mathfrak{H}_0$, defined over $\tilde{\mathfrak{R}}$, such that $\tilde{\varrho}_H$ is contained in ϱ_{K}.*

Our problem is thus reduced to seeking conditions such that $\tilde{\varrho}_H$ is contained in ϱ_{K} for some $\mathrm{K} \in \mathfrak{H} \ominus \mathfrak{H}_0$. Let $\boldsymbol{\lambda}(R, \boldsymbol{x})$ be an $N(n+1) \times N(n+1)$ element nonsingular matrix all of whose elements are of class C^1 for all $\boldsymbol{x}$ in $\mathscr{D}_n^*$ and for all $\{\tilde{R}\}$ in $\tilde{\mathfrak{R}}$. For each such matrix we may define a new operator

$$\{\mathfrak{O}_{\mathrm{K}\lambda}\} \underset{\tilde{\mathfrak{R}}}{\overset{\text{def}}{=}} \boldsymbol{\lambda}\{\mathfrak{O}_{\mathrm{K}}\} \tag{2.32}$$

where the indicated multiplication is matrix multiplication. Under the continuity assumption on $\boldsymbol{\lambda}$ we may interpret $\{\mathfrak{O}_{K\lambda}\}$ as a map of $\tilde{\mathfrak{R}}$ into $\mathfrak{S}$

$$\{\mathfrak{O}_{K\lambda}\}: \tilde{\mathfrak{R}} \xrightarrow{\text{onto}} \mathfrak{S}_{K\lambda} \subset \mathfrak{S}. \tag{2.33}$$

Combining equations (2.30) and (2.32), we have

$$\{\mathfrak{O}_K\}(r) = \{s\} \in \mathfrak{S}_K \in \mathfrak{S}, \qquad \{r\} \in \tilde{\mathfrak{R}},$$
$$\{\mathfrak{O}_{K\lambda}\}(r) = \boldsymbol{\lambda}\{\mathfrak{O}_K\}(r) = \boldsymbol{\lambda}\{s\} \in \mathfrak{S}_{K\lambda} \subset \mathfrak{S},$$

and hence the matrix multiplication of $\{\mathfrak{O}_K\}$ by $\boldsymbol{\lambda}$ may be considered to induce a transformation of $\mathfrak{S}_K$ to $\mathfrak{S}_{K\lambda}$ as defined by

$$\boldsymbol{\lambda}\{s\} = \{s_\lambda\}; \qquad \{s\} \in \mathfrak{S}_K; \qquad \{s_\lambda\} \in \mathfrak{S}_{K\lambda}$$
$$\mathfrak{S}_{K\lambda} = \bigcup_{\mathfrak{S}_K} \{s_\lambda\}.$$

We shall refer to such a transformation induced on $\mathfrak{S}_K$ by equations (2.32) and (2.33) as a **lambda transformation** and the matrix $\boldsymbol{\lambda}$ as a **lambda matrix**. The collection of all lambda matrices we shall denote as $\boldsymbol{\lambda}$.

Let $\mathfrak{M}_{K\lambda}$ be the set of all elements $\{m\}$ of $\tilde{\mathfrak{R}}$ which satisfy

$$\{\tilde{\mathfrak{O}}_H\}(m) = \{s\} = \boldsymbol{\lambda}\{\mathfrak{O}_K\}(m) = \{\mathfrak{O}_{K\lambda}\}(m)$$

for $\boldsymbol{\lambda}$ an element of $\boldsymbol{\lambda}$ and K an element of $\mathfrak{H} \ominus \mathfrak{H}_0$. For simplicity, we shall denote the set $\mathfrak{M}_{K\lambda}$ and its defining relation by

$$\{\tilde{\mathfrak{O}}_H\} \underset{\mathfrak{M}_{K\lambda}}{=} \{\mathfrak{O}_{K\lambda}\}. \tag{2.34}$$

By equation (2.28a) we have

$$\boldsymbol{M}^{-1\prime}\{\mathfrak{O}_H^*\} \underset{\tilde{\mathfrak{R}}}{=} \{\tilde{\mathfrak{O}}_H\},$$

and hence equation (2.34) is equivalent to

$$\boldsymbol{M}^{-1\prime}\{\mathfrak{O}_H^*\} \underset{\mathfrak{M}_{K\lambda}}{=} \boldsymbol{\lambda}\{\mathfrak{O}_K\},$$

since $\mathfrak{M}_{K\lambda}$ is contained in $\tilde{\mathfrak{R}}$. Noting that $\boldsymbol{M}$ is nonsingular by definition, we have

$$\{\mathfrak{O}_H^*\} \underset{\mathfrak{M}_{K\lambda}}{=} \boldsymbol{M}'\boldsymbol{\lambda}\{\mathfrak{O}_K\}.$$

Thus, we consider the set $\mathfrak{M}^*_{K\lambda}$ defined by

$$\{\mathfrak{O}_H^*\} \underset{\mathfrak{M}^*_{K\lambda}}{=} \boldsymbol{\lambda}\{\mathfrak{O}_K\}. \tag{2.35}$$

Let $\{u\}$ be an element of $\tilde{\varrho}_H$ and in addition an element of $\mathfrak{M}^*_{K\lambda}$ for some $\boldsymbol{\lambda} \in \boldsymbol{\lambda}$, say $\boldsymbol{\lambda}_0$. The assumption that $\{u\}$ is an element of $\tilde{\varrho}_H$ implies that $\{\mathfrak{O}_H^*\}(u) = \{\Phi\}$; and $\{u\}$ an element of $\mathfrak{M}^*_{K\lambda_0}$ implies, by equation (2.35), that $\{\Phi\} = \boldsymbol{\lambda}_0\{\mathfrak{O}_K\}(u)$. If $\{s_K\}$ is the image, in $\mathfrak{S}$, of $\{u\}$ under $\{\mathfrak{O}_K\}$ so that $\{\mathfrak{O}_K\}(u) = \{s_K\}$, then the consideration immediately preceding implies that $\{\Phi\} = \boldsymbol{\lambda}\{s_K\}$ and hence that $\{u\}$ is an element of ϱ_K, since $\boldsymbol{\lambda}_0$ is nonsingular by hypothesis. Thus, a sufficient condition that $\tilde{\varrho}_H$ be contained in ϱ_K is that there exists at least one lambda matrix, say λ_0, such that $\tilde{\varrho}_H$ is contained in $\mathfrak{M}^*_{K\lambda_0}$. Conversely, if $\tilde{\varrho}_H$ is contained in ϱ_K, then for any $\{u\}$ an element of $\tilde{\varrho}_H$ we have $\{\mathfrak{O}_H^*\}(u) = \{\Phi\}$ and $\{\mathfrak{O}_K\}(u) = \{\Phi\}$.

Hence, equation (2.35) is satisfied for $\{u\} \in \tilde{\varrho}_H$ and for any $\boldsymbol{\lambda} \in \boldsymbol{\lambda}$ so that $\tilde{\varrho}_H$ is contained in $\mathfrak{M}^*_{K\lambda}$. We have thus proved

*Lemma 2.2. *A necessary and sufficient condition for $\tilde{\varrho}_H$ to be contained in ϱ_K is that there exist at least one lambda matrix $\boldsymbol{\lambda}_0 \in \boldsymbol{\lambda}$ such that the set $\mathfrak{M}^*_{K\lambda_0}$, defined by $\{\mathfrak{O}^*_H\} \underset{\mathfrak{M}_{K\lambda_0}}{=\!=} \boldsymbol{\lambda}_0 \{\mathfrak{O}_K\}$, contains $\tilde{\varrho}_H$.*

Combining the above two lemmas, we have

*Lemma 2.3. *A necessary and sufficient condition that a map T of the collection $\mathcal{T}$ map all points of $\mathfrak{R}$ which form Hamiltonian systems base H, for H a fixed element of $\mathfrak{H} \ominus \mathfrak{H}_0$, into Hamiltonian systems in $\widetilde{\mathfrak{R}}$ is that there exist at least one function $K \in \mathfrak{H} \ominus \mathfrak{H}_0$, defined over $\widetilde{\mathfrak{R}}$, and at least one lambda matrix $\boldsymbol{\lambda}_0$ such that the set of points $\mathfrak{M}^*_{K\lambda}$, defined by*

$$\{\mathfrak{O}^*_H\} \underset{\mathfrak{M}_{K\lambda_0}}{=\!=} \boldsymbol{\lambda}_0 \{\mathfrak{O}_K\}, \tag{2.35}$$

shall contain $\tilde{\varrho}_H$. If these conditions are satisfied, then K *is the new base function corresponding to H under T.*

This lemma is the basis for the investigation of the conditions which must be placed on a map T in order for it to be canonical. Using equations (2.28b), (2.29) and (2.35) and the above lemma, we see that a necessary and sufficient condition that T map Hamiltonian systems base H, for fixed H, into Hamiltonian systems is that there exist at least one function K and at least one lambda matrix such that

$$\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} \bullet (\vec{V}\{\widetilde{R}\} - \underset{0}{\vec{V}}\{\widetilde{R}\}) + \{H_{,\widetilde{R}}\} = \boldsymbol{\lambda}_0 (\vec{\boldsymbol{I}} \bullet \vec{V}\{\widetilde{R}\} + \{K_{,\widetilde{R}}\}) \tag{2.36}$$

shall hold for all points $\mathfrak{M}^*_{K\lambda_0}$ which are elements of $\widetilde{\mathfrak{R}}$ where $\mathfrak{M}^*_{K\lambda}$ is such that it contains $\tilde{\varrho}_H$. At this point we must distinguish between two cases.

Case 1. The point set $\tilde{\varrho}_H$ forms at least a one parameter continuum. In this case, the requirement that equation (2.36) holds for all points of a set $\mathfrak{M}^*_{K\lambda_0} \subset \mathfrak{R}$ and that $\mathfrak{M}^*_{K\lambda_0}$ contains $\tilde{\varrho}_H$ implies that we require equation (2.36) to be satisfied over a continuum of points of $\widetilde{\mathfrak{R}}$. This can only be the case if the coefficients of the various derivatives of $\{\widetilde{R}\}$ in equation (2.36) vanish separately.

Case 2. The point set $\tilde{\varrho}_H$ is composed of a finite number of elements. In this case we proceed as follows. Since it is required that equation (2.36) hold for all points of a set $\mathfrak{M}^*_{K\lambda_0}$ which contains $\tilde{\varrho}_H$, we require $\mathfrak{M}^*_{K\lambda_0}$ to be at least a one-parameter continuum containing $\tilde{\varrho}_H$. For this choice, equation (2.36) can only hold over such a continuum if the coefficients of the various derivatives of $\{\widetilde{R}\}$ vanish separately. It must be noted that in this case we are actually placing added conditions on the problem since there can exist functions K for which equation (2.36) is only satisfied for a set of points of $\widetilde{\mathfrak{R}}$ which are finite in number and contain $\tilde{\varrho}_H$. Thus, in this case, the following analysis will only result in sufficient conditions rather than both necessary and sufficient conditions.

The requirement that the coefficients of the various derivatives of $\{\widetilde{R}\}$ vanish separately results in

$$\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} = \boldsymbol{\lambda}\vec{\boldsymbol{I}}, \tag{2.37}$$

$$\boldsymbol{\lambda}\{K_{,\widetilde{R}}\} = \{H_{,\widetilde{R}}\} - \boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} \bullet \underset{0}{\vec{V}}\{\widetilde{R}\}, \tag{2.38}$$

since $\underset{0}{\vec{\nabla}}$ is not a differential operator but assigns a function of $(\tilde{R}, \boldsymbol{x})$ as seen from its definition (equation (2.20)). Using equation (2.37) in equation (2.38) gives, upon multiplication on the left by $\boldsymbol{\lambda}^{-1}$,

$$\{\mathrm{K}_{,\tilde{R}}\} = \boldsymbol{\lambda}^{-1}\{H_{,\tilde{R}}\} - \vec{\boldsymbol{I}} \bullet \underset{0}{\vec{\nabla}}\{\tilde{R}\}. \tag{2.39}$$

Examining equation (2.39), we see that it is a first order differential equation for the determination of K.

A necessary and sufficient condition for the existence of a general solution to equation (2.39) is that the matrix $\boldsymbol{J} = \{\mathrm{K}_{,\tilde{R}}\}_{,\{\tilde{R}\}}$ be symmetric, since all terms in equation (2.39) are of class C^1. Setting $\boldsymbol{\lambda}^{-1} = \boldsymbol{\eta}$, we have, since $\vec{\boldsymbol{I}}$ is a constant vectrix,

$$\boldsymbol{J} = (\boldsymbol{\eta}\{H_{,\tilde{R}}\})_{,\{\tilde{R}\}} - \vec{\boldsymbol{I}} \bullet (\underset{0}{\vec{\nabla}}\{\tilde{R}\})_{,\{\tilde{R}\}}.$$

But

$$(\underset{0}{\vec{\nabla}}\{\tilde{R}\})_{,\{\tilde{R}\}} = (\underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1})\,\boldsymbol{M},$$

and hence

$$\boldsymbol{J} = (\boldsymbol{\eta}\{H_{,\tilde{R}}\})_{,\{\tilde{R}\}} - (\vec{\boldsymbol{I}} \bullet \underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1})\,\boldsymbol{M}.$$

The integrability conditions of equation (2.39) thus read:

$$\boldsymbol{O} = (\boldsymbol{\eta}\{H_{,\tilde{R}}\})_{,\{\tilde{R}\}} - [(\boldsymbol{\eta}\{H_{,\tilde{R}}\})_{,\{\tilde{R}\}}]' - (\vec{\boldsymbol{I}} \bullet \underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1})\,\boldsymbol{M} + \boldsymbol{M}'\,\underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1\prime} \bullet \vec{\boldsymbol{I}}'$$

which, since $\vec{\boldsymbol{I}}' = -\vec{\boldsymbol{I}}$, gives

$$\boldsymbol{O} = (\boldsymbol{\eta}\{H_{,\tilde{R}}\})_{,\{\tilde{R}\}} - [(\boldsymbol{\eta}\{H_{,\tilde{R}}\})_{,\{\tilde{R}\}}]' - (\vec{\boldsymbol{I}} \bullet \underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1})\boldsymbol{M} - \boldsymbol{M}'\,\underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1\prime} \bullet \vec{\boldsymbol{I}}.$$

Set

$$\boldsymbol{v} = -(\vec{\boldsymbol{I}} \bullet \underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1})\,\boldsymbol{M} - \boldsymbol{M}'\,\underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1\prime} \bullet \vec{\boldsymbol{I}};$$

then

$$\begin{aligned}\boldsymbol{M}^{-1\prime}\,\boldsymbol{v}\,\boldsymbol{M}^{-1} &= -\boldsymbol{M}^{-1\prime}\,\vec{\boldsymbol{I}} \bullet \underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1} - (\underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1})' \bullet \vec{\boldsymbol{I}}\boldsymbol{M}^{-1},\\ &= -\underset{0}{\vec{\nabla}} \bullet (\boldsymbol{M}^{-1\prime}\,\vec{\boldsymbol{I}}\boldsymbol{M}^{-1}).\end{aligned}$$

From equation (2.37) we have $\boldsymbol{\eta}\boldsymbol{M}'\vec{\boldsymbol{I}} = \vec{\boldsymbol{I}}\boldsymbol{M}^{-1}$ where $\boldsymbol{\eta} = \boldsymbol{\lambda}^{-1}$, and hence

$$\boldsymbol{M}^{-1\prime}\vec{\boldsymbol{I}}\boldsymbol{M}^{-1} = \boldsymbol{M}^{-1\prime}\boldsymbol{\eta}\boldsymbol{M}'\vec{\boldsymbol{I}}$$

so that

$$\boldsymbol{v} = -\boldsymbol{M}'\,[\underset{0}{\vec{\nabla}}(\boldsymbol{M}^{-1\prime}\boldsymbol{\eta}\boldsymbol{M}')] \bullet \vec{\boldsymbol{I}}\boldsymbol{M}.$$

Thus,

$$\boldsymbol{O} = (\boldsymbol{\eta}\{H_{,\tilde{R}}\})_{,\{\tilde{R}\}} - [(\boldsymbol{\eta}\{H_{,\tilde{R}}\})_{,\{\tilde{R}\}}]' - \boldsymbol{M}'\,[\underset{0}{\vec{\nabla}}(\boldsymbol{M}^{-1\prime}\boldsymbol{\eta}\boldsymbol{M}')] \bullet \vec{\boldsymbol{I}}\boldsymbol{M} \tag{2.40}$$

are the integrability conditions for equation (2.39). Since H and $\boldsymbol{M}$ are given (that is, $\boldsymbol{M}$ satisfies equation (2.37)), equation (2.40) is a partial differential equation for the determination of $\boldsymbol{\eta}(\tilde{R}, \boldsymbol{x})$. We have thus proved

***Theorem 2.1.** *A necessary and sufficient condition in the case in which* $\tilde{\varrho}_H$ *forms a continuum, and a sufficient condition in the case in which* $\tilde{\varrho}_H$ *is composed of only a finite number of elements, for a map* T *of the collection* $\mathscr{T}$ *to map a Hamiltonian systems base* H*, for fixed* $H \in \mathfrak{H} \ominus \mathfrak{H}_0$*, into a Hamiltonian system, is that* (1) *the Jacobian matrix of the transformation satisfies equation* (2.37) *and* (2) *that there exists a solution* $\boldsymbol{\eta}(\tilde{R}, \boldsymbol{x})$ *to equation* (2.40). *In addition, if the above conditions are satisfied, then equation* (2.39) *is completely integrable and determines the new basis function* K.

By definition, a map of the collection $\mathscr{T}$ is canonical if and only if it maps all Hamiltonian systems formed from all bases functions into Hamiltonian systems; or equivalently, it maps all Hamiltonian systems formed from all bases functions of the collection $\mathfrak{H} \ominus \mathfrak{H}_0$ into Hamiltonian systems. In order to apply the results of Theorem 2.1 to characterize canonical maps, we must distinguish between those Hamiltonian systems whose solutions form at least a one-parameter continuum and those Hamiltonian systems which admit only a finite number of solutions.

Denote by $\mathfrak{H}_1$ the sub-collection of $\mathfrak{H} \ominus \mathfrak{H}_0$ which are such that the Hamiltonian systems formed with basis functions which are members of this sub-collection admit at least a one parameter continuum of solutions, and by $\mathfrak{H}_2$ the sub-collection of $\mathfrak{H} \ominus \mathfrak{H}_0$ for which the Hamiltonian systems formed from this sub-collection of basis functions admit only a finite number of solutions. Since all Hamiltonian systems formed from basis functions of the collection $\mathfrak{H} \ominus \mathfrak{H}_0$ are such that they admit at least one solution and the conditions defining $\mathfrak{H}_1$ and $\mathfrak{H}_2$ are mutually exclusive possibilities, we see that the set-theoretic union of $\mathfrak{H}_1$ and $\mathfrak{H}_2$ contains and is contained in $\mathfrak{H} \ominus \mathfrak{H}_0$ and that the intersection of $\mathfrak{H}_1$ and $\mathfrak{H}_2$ is empty. Let $\mathscr{T}_1$ be the collection of all maps T of the collection $\mathscr{T}$ which map Hamiltonian systems formed from basis functions of the sub-collection $\mathfrak{H}_1$ into Hamiltonian systems, $\mathscr{T}_2$ the collection of all maps T of the collection $\mathscr{T}$ which map Hamiltonian systems formed from basis functions of the sub-collection $\mathfrak{H}_2$ into Hamiltonian systems, and $\mathscr{T}_c$ the collection of all canonical maps. From the definition of a canonical map, we see that $\mathscr{T}_c$ is equal to the intersection of $\mathscr{T}_1$ and $\mathscr{T}_2$. Since Theorem 2.1 establishes necessary and sufficient conditions for a map to be an element of the collection $\mathscr{T}_1$, and a sufficient condition for a map to be of the collection $\mathscr{T}_2$, for given H; and since $\mathscr{T}_c$ is the intersection of $\mathscr{T}_1$ and $\mathscr{T}_2$; we see that a map T is canonical if and only if conditions (1) and (2) of Theorem 2.1 hold for all H contained in $\mathfrak{H} \ominus \mathfrak{H}_0$.

Examining condition (1) of Theorem 2.1, we see that $\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} = \boldsymbol{\lambda}\vec{\boldsymbol{I}}$ must be satisfied for all H which are elements of $\mathfrak{H} \ominus \mathfrak{H}_0$ simultaneously. If T is to map all Hamiltonian systems simultaneously into Hamiltonian systems, then it must be independent of the base function for any particular system, and hence the Jacobian matrix of the transformation must be independent of H. This can only be the case if $\boldsymbol{\lambda}$ is independent of H, since we require the associated Jacobian matrix of a canonical map to satisfy equation (2.48).

Condition (2) of Theorem 2.1 requires that there exist at least one lambda matrix $\boldsymbol{\lambda} = \boldsymbol{\eta}^{-1}$ which satisfies equation (2.40). Since equation (2.40) must be satisfied simultaneously for all H which are elements of $\mathfrak{H} \ominus \mathfrak{H}_0$, and in addition

it is required that $\boldsymbol{\lambda}$ must be independent of H, equation (2.40) must be an identity in H. This can be true only if the coefficients of H and its derivatives vanish separately in equation (2.40). Thus

$$\boldsymbol{\lambda}^{-1}\{H_{,\tilde{R}}\}_{,\{\tilde{R}\}} - \{H_{,\tilde{R}}\}_{,\{\tilde{R}\}}\boldsymbol{\lambda}^{-1\prime} = \boldsymbol{O}, \tag{2.41}$$

$$\boldsymbol{\lambda}^{-1}_{,\{\tilde{R}\}}\{H_{,R}\} - \{H_{,\tilde{R}}\}'(\boldsymbol{\lambda}^{-1}_{,\{\tilde{R}\}})' = \boldsymbol{O}, \tag{2.42}$$

$$\boldsymbol{M}'[\underset{0}{\vec{\nabla}}(\boldsymbol{M}^{-1\prime}\boldsymbol{\lambda}\boldsymbol{M}')]\bullet\vec{\boldsymbol{I}}\boldsymbol{M} = \boldsymbol{O} \tag{2.43}$$

for all H contained in $\mathfrak{H}\ominus\mathfrak{H}_0$. Since equation (2.41) must hold for all $H\in\mathfrak{H}\ominus\mathfrak{H}_0$, then it must be satisfied by an H such that $\{H_{,\tilde{R}}\}_{,\{\tilde{R}\}} = \boldsymbol{E}^*$. For this case we have $\boldsymbol{\lambda}^{-1} - \boldsymbol{\lambda}^{-1\prime} = \boldsymbol{O}$ so that $\boldsymbol{\lambda}^{-1}$ must be symmetric. Rewriting equation (2.41) to reflect this result, we obtain

$$\boldsymbol{\lambda}^{-1}\{H_{,\tilde{R}}\}_{,\{\tilde{R}\}} - \{H_{,\tilde{R}}\}_{,\{\tilde{R}\}}\boldsymbol{\lambda}^{-1} = \boldsymbol{O}.$$

Since the only matrix which commutes with all matrices is a multiple of the identity matrix, we may conclude that $\boldsymbol{\lambda}^{-1}$ must have the general form $\boldsymbol{\lambda}^{-1} = \mu(\tilde{R},\boldsymbol{x})\,\boldsymbol{E}^*$, where μ is a scalar function. Substituting the form of $\boldsymbol{\lambda}^{-1}$ thus far obtained into equation (2.42) gives

$$\{\mu_{,\tilde{R}}\}\{H_{,\tilde{R}}\}' - \{H_{,\tilde{R}}\}\{\mu_{,\tilde{R}}\}' = \boldsymbol{O},$$

which can only be satisfied for all $H\in\mathfrak{H}\ominus\mathfrak{H}_0$ and for $\boldsymbol{\lambda}$ independent of H by $\{\mu_{,\tilde{R}}\} = \{0\}$. Thus $\boldsymbol{\lambda}^{-1}$ must have the form $\boldsymbol{\lambda}^{-1} = \mu(\boldsymbol{x})\,\boldsymbol{E}^*$. Substituting into equation (2.43) gives

$$\boldsymbol{M}'[\underset{0}{\vec{\nabla}}\mu(\boldsymbol{x})]\bullet\vec{\boldsymbol{I}}\boldsymbol{M} = \boldsymbol{O}.$$

Since $\boldsymbol{M}$ is nonsingular by hypothesis, we have

$$\underset{0}{\vec{\nabla}}\mu(\boldsymbol{x})\bullet\vec{\boldsymbol{I}} = \boldsymbol{O}.$$

Thus $\boldsymbol{\lambda}^{-1}$ is a constant multiple of the identity matrix. Since lambda matrices are nonsingular by definition, the constant multiplying the identity matrix must be non-zero. Thus $\boldsymbol{\lambda}^{-1} = \mu\boldsymbol{E}^*$, and hence $\boldsymbol{\lambda} = \frac{1}{\mu}\boldsymbol{E}^*$ where μ is a non-zero constant.

From the above considerations we see that if there exists a non-zero constant μ such that equation (2.37) is satisfied where $\boldsymbol{\lambda} = \frac{1}{\mu}\boldsymbol{E}$, then equation (2.38) is integrable and these conditions hold for all $H\in\mathfrak{H}\ominus\mathfrak{H}_0$.

Summarizing, we have

$$\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} = \frac{1}{\mu}\vec{\boldsymbol{I}} \tag{2.44}$$

and

$$\{\mathrm{K}_{,\tilde{R}}\} = \mu\{H_{,\tilde{R}}\} - \vec{\boldsymbol{I}}\bullet\underset{0}{\vec{\nabla}}\{\tilde{R}\}. \tag{2.45}$$

Since equation (2.40) is satisfied by $\boldsymbol{\lambda} = \frac{1}{\mu}\boldsymbol{E}$, there exists a function $U(\tilde{R},\boldsymbol{x})$ of class C^2 in its arguments (that is, $\mathscr{T}$ was assumed to be of class C^2 in $\boldsymbol{x}$) such that

$$\{U_{,\tilde{R}}\} = \vec{\boldsymbol{I}}\bullet\underset{0}{\vec{\nabla}}\{\tilde{R}\}. \tag{2.46}$$

Substituting equation (2.46) into equation (2.45) and integrating gives

$$\mathrm{K} = \mu H - U \tag{2.47}$$

to within an arbitrary additive function of $\boldsymbol{x}$ alone. It is evident that

$$\{\mathrm{K}_{,\tilde{R}}\} = \{(\mathrm{K} + v(\boldsymbol{x}))_{,\tilde{R}}\}$$

for all $\{\widetilde{R}\}$ since $\{v(\boldsymbol{x})_{,\tilde{R}}\} = O$. We may thus consider two base functions as equivalent if they differ by a function of $\boldsymbol{x}$ alone. Hence, we may take equation (2.47) as the general solution of equation (2.45). The last condition, namely that K be an element of $\mathfrak{H} \ominus \mathfrak{H}_0$, is obviously fulfilled by inspection of equation (2.47), since both H and U are elements of $\mathfrak{H} \ominus \mathfrak{H}_0$.

We have thus proved the following fundamental

***Theorem 2.2.** *A map T of the collection $\mathscr{T}$ is a canonical map if and only if there exists a nonzero constant μ such that $\boldsymbol{M}$, the associated Jacobian matrix of T, satisfies the vectrix equation*

$$\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} = \frac{1}{\mu}\vec{\boldsymbol{I}}, \tag{2.48}$$

in which case the new basis function, K, *is given by*

$$\mathrm{K} = \mu H - U \tag{2.49}$$

where

$$\{U_{,\tilde{R}}\} = \vec{\boldsymbol{I}} \bullet \underset{0}{\vec{\nabla}}\{\widetilde{R}\}. \tag{2.50}$$

It is of interest to note that this theorem for the case $n=1$ is identical with the results given by SIEGEL[4] and is equivalent to the results of WINTNER[5], as will be shown in Section IV.

Generalized Symplectic Matrices

A matrix $\boldsymbol{M}$ which satisfies equation (2.48) in the case $n=1$ is termed a symplectic[6] matrix. A matrix $\boldsymbol{M}$ which satisfies equation (2.48) for $n > 1$ will be referred to as a **generalized symplectic matrix of multiplier** μ, the constant μ being referred to as the **multiplier** of the canonical map. The function $U(P, Q, \boldsymbol{x})$, defined by equation (2.50), will be referred to as the **remainder function**. From the form of equations (2.50) and (2.49) it is evident that if the transformation equations do not involve the independent variables explicitly, the remainder function is, at most, a function of $\boldsymbol{x}$ alone and may thus be deleted with no loss of generality. Thus, in this case equation (2.49) states that the new basis function differs from the old basis function under the transformation T by a nonzero, constant multiple of

$$H\big(p(P, Q, \boldsymbol{x}), \mathrm{q}(P, Q, \boldsymbol{x}), \boldsymbol{x}\big).$$

The collection of all canonical maps may be partially partitioned into several natural subcollections. A canonical map is said to be **uniform** if the corresponding transformation equations do not involve the independent variables explicitly, or

[4] SIEGEL, C. L.: Vorlesungen über Himmelsmechanik. Berlin-Göttingen-Heidelberg: Springer 1956.

[5] WINTNER, A.: The Analytical Foundations of Celestical Mechanics, equation 23, p. 11. Princeton University Press 1956.

[6] SIEGEL, C. L.: Vorlesungen über Himmelsmechanik, p. 8.

equivalently, the remainder function may be set equal to zero with no loss of generality. A canonical map is said to be **basic** if its multiplier is unity and is said to be **absolutely uniform** if it is both basic and uniform.

To conclude the analysis of this section, we state the following corollary to Theorem 2.1. This corollary is immediate from the definition of a generalized symplectic matrix of multiplier μ.

* **Corollary.** A map T of the collection $\mathscr{T}$ is a canonical map if and only if there exists a scalar μ such that the Jacobian matrix is a generalized symplectic matrix of multiplier μ.

The definition of a canonical map given here differs from those presented in a majority of the literature. WHITTAKER[7] considers only canonical maps in which $\mu=1$ and defines this subcollection by requiring that the differential form $P_{\alpha 1} d Q_\alpha - p_{\alpha 1} d q_2$ form a total differential for the case of one independent variable, thus defining canonical maps in terms of a first PFAFF'S[8] expression. It was JACOBI[9] who first proved that canonical maps, as defined in the manner of WHITTAKER, are such as to leave the form of the Hamiltonian equations invariant.

Section III. An Alternate Development

In this section, we take up an alternate development of the conditions which must be placed on a map T of the collection $\mathscr{T}$ in order that it shall be canonical. This alternate development is included for two reasons. First, it is based on the Lagrangean formulation and thus will provide more direct application of the results to physical problems; and second, it considers the problem from the variational viewpoint.

Variational Considerations

Let α range over the finite index set $\mathscr{A}$ of numerosity N and let $f(\mathrm{q}_\alpha; \mathrm{q}_{\alpha,i}; \boldsymbol{x})$ be a real function of class C^2 in its $N(n+1)+n$ arguments q_α; $\mathrm{q}_{\alpha,i}$, $\boldsymbol{x}$. In addition, let $\mathrm{q}_\alpha(\boldsymbol{x})$ be of class C^2 in some given $\mathscr{D}_n^*$, which assume given values on the boundary of $\mathscr{D}_n^*$. We wish to establish conditions on the $\mathrm{q}_\alpha(\boldsymbol{x})$ such that the functional

$$\mathscr{I} = \int_{\mathscr{D}_n^*} f\left(\mathrm{q}_\alpha(\boldsymbol{x}); \mathrm{q}_{\alpha,i}(\boldsymbol{x}); \boldsymbol{x}\right) d v_n$$

shall have an extremal value relative to the choice of the $\mathrm{q}_\alpha(\boldsymbol{x})$, for the functional form of f fixed.

In order to obtain the required conditions, we consider a parametric imbedding of the q_α as follows:

$$q_\alpha(\eta, \boldsymbol{x}), \qquad -1<\eta<1,$$

$$q_\alpha(\eta, \boldsymbol{x})_{,\eta} \in C^1, \qquad q_\alpha(O, \boldsymbol{x}) = \mathrm{q}_\alpha(\boldsymbol{x})$$

where the last equation is interpreted as stating that the choice of q_α which extremized $\mathscr{I}$ is given by $\mathrm{q}_\alpha(\boldsymbol{x})$. Since q_α assume given values on the boundary, we require in addition that

$$q_\alpha(\eta, \boldsymbol{x})|_{\mathscr{D}_n^* \ominus \mathscr{D}_n} = \mathrm{q}_\alpha(\boldsymbol{x})|_{\mathscr{D}_n^* \ominus \mathscr{D}_n}.$$

[7] WHITTAKER, E. T.: A Treatise on the Analytical Dynamics of Particles and Rigid Bodies, Section 126.

[8] PFAFF, J. F.: Abh. d. Berl. Acad. 76 (1814—1815).

[9] JACOBI, K. G. J.: Compt. Rendus 61 (1837).

Under the above parametric imbedding, the functional $\mathscr{I}$ becomes a function of the parameter η (that is,

$$\mathscr{I}(\eta) = \int_{\mathscr{D}_n^*} f\left(q_\alpha(\eta, \boldsymbol{x}); q_{\alpha,i}(\eta, \boldsymbol{x}); \boldsymbol{x}\right) dv_n\right),$$

so that $\mathscr{I}(\eta)$ has an extremum for the value $\eta = 0$ under the condition $d\mathscr{I}(\eta)/d\eta = 0$. Let $\delta h(\eta, \boldsymbol{x})$ denote $\partial/\partial\eta\,(h(\eta, \boldsymbol{x}))$, so that

$$\delta\mathscr{I}(\eta) = \int_{\mathscr{D}_n^*} \{f_{,q_\alpha}\,\delta q_\alpha + f_{,q_{\alpha,i}}\,\delta q_{\alpha,i}\}\,dv_n.$$

Consider the vectrix $\vec{J}$ defined by

$$\vec{J} \stackrel{\text{def}}{=} \vec{e}_i f_{,q_{\alpha,i}}\,\delta q_\alpha. \tag{3.1}$$

Now

$$\int_{\mathscr{D}_n^*} \vec{\nabla}\bullet\vec{J}\,dv_n = \int_{\mathscr{D}_n^* \ominus \mathscr{D}_n} \vec{J}\bullet\vec{N}\,dS = \int_{\mathscr{D}_n^* \ominus \mathscr{D}_n} f_{,q_{\alpha,i}}\,\delta q_\alpha N_i\,dS$$

by equation (1.16). However, δq_α vanishes on $\mathscr{D}_n^* \ominus \mathscr{D}_n$ since q_α is independent of η on $\mathscr{D}_n^* \ominus \mathscr{D}_n$ by hypothesis. Hence

$$0 = \int_{\mathscr{D}_n^*} \vec{\nabla}\bullet\vec{J}\,dv_n = \int_{\mathscr{D}_n^*} \{(f_{,q_{\alpha,i}})_{,i} + f_{,q_{\alpha,i}}(\delta q_\alpha)_{,i}\}\,dv_n.$$

From the definition of the operation δ, it is evident that $\delta[(\ )_{,i}] = [\delta(\ )]_{,i}$ under the assumed continuity conditions, and so

$$\int_{\mathscr{D}_n^*} f_{,q_{\alpha,i}}\,\delta q_{\alpha,i}\,dv_n = -\int_{\mathscr{D}_n^*} (f_{,q_{\alpha,i}})_{,i}\,\delta q_\alpha\,dv_n.$$

Substituting this result into the expression for $\delta\mathscr{I}(\eta)$, we obtain

$$\delta\mathscr{I}(\eta) = \int_{\mathscr{D}_n^*} \{f_{,q_\alpha} - (f_{,q_{\alpha,i}})_{,i}\}\,\delta q_\alpha\,dv_n,$$

or equivalently

$$\delta\mathscr{I}(\eta) = \int_{\mathscr{D}_n^*} \{E|f\}_{q_\alpha}\,\delta q_\alpha\,dv_n$$

where

$$\{E|f\}_{q_\alpha} \stackrel{\text{def}}{=} f_{,q_\alpha} - (f_{,q_{\alpha,i}})_{,i} \tag{3.2}$$

is the Euler-Lagrange operator. Now, $\mathscr{I}(\eta)$ will be stationary if $\delta\mathscr{I}(\eta)|_{\eta=0} = 0$. Since $\delta q_\alpha|_{\eta=0}$ is arbitrary, we obtain as a necessary condition for $\mathscr{I}$ to be stationary

$$\{E|f\}_{q_\alpha} = 0. \tag{3.3}$$

That this condition is sufficient is immediate. Note, from equations (3.1) and (3.2), that

$$\{E|f\}_{q_\alpha}\,\delta q_\alpha = \delta f - \vec{\nabla}\bullet\vec{J}. \tag{3.2a}$$

Invariance of the Euler-Lagrange Operator

Consider now an arbitrary transformation T on q_α of class $C^{[2]}$,

$$q_\alpha = q_\alpha(Q_\beta, \boldsymbol{x}),$$

and imbedded Q_β in a parametrization by the equations

$$Q_\beta(\eta, \boldsymbol{x}) = Q_\beta\left(q_\alpha(\eta, \boldsymbol{x}), \boldsymbol{x}\right).$$

Now

$$q_{\alpha,i} = \partial_i q_\alpha + q_{\alpha,Q_\beta} Q_{\beta,i}$$

so that

$$q_{\alpha,Q_{\beta,i}} = 0, \qquad q_{\alpha,i,Q_{\beta,j}} = \delta_{ij} q_{\alpha,Q_\beta} \tag{3.4}$$

and

$$\delta q_\alpha = q_{\alpha,Q_\beta} \delta Q_\beta. \tag{3.5}$$

Under T we have

$$f_{,Q_{\beta,j}} = f_{,q_\alpha} q_{\alpha,Q_{\beta,j}} + f_{,q_{\alpha,i}} q_{\alpha,i,Q_{\beta,j}},$$

which by equations (3.4) gives

$$f_{,Q_{\beta,j}} = f_{,q_{\alpha,j}} q_{\alpha,Q_\beta}.$$

Hence

$$f_{,Q_{\beta,j}} \delta Q_\beta = f_{,q_{\alpha,j}} q_{\alpha,Q_\beta} \delta Q_\beta,$$

which by equations (3.5) gives

$$f_{,Q_{\beta,j}} \delta Q_\beta = f_{,q_{\alpha,j}} \delta q_\alpha = J_j$$

so that $\vec{J}$ is invariant under T. Now from equation (3.2a) we have

$$\delta f - \vec{\nabla} \bullet \vec{J} = \{E|f\}_{q_\alpha} \delta q_\alpha,$$

which, since both δf and $\vec{J}$ are invariant under T, gives

$$\delta f - \vec{\nabla} \bullet \vec{J} = \{E|f\}_{Q_\beta} \delta Q_\beta,$$

and hence, by equation (3.5),

$$\{E|f\}_{Q_\beta} \delta Q_\beta = \{E|f\}_{q_\alpha} \delta q_\alpha = \{E|f\}_{q_\alpha} q_{\alpha,Q_\beta} \delta Q_\beta$$

so that

$$\{E|f\}_{Q_\beta} = \{E|f\}_{q_\alpha} q_{\alpha,Q_\beta} \tag{3.6}$$

under any T of class $C^{[2]}$. Thus we have the well known

Theorem 3.1. *The vanishing of the Euler-Lagrange operator is a necessary and sufficient condition for $\mathscr{I}(\eta)$ to be extremal under all transformations T in its functional arguments of class $C^{[2]}$.*

It is evident from the definition of the Euler-Lagrange operator that it is a linear operator on a collection of functions of the same nature as the f above, that is,

$$\{E| \sum_\mu \lambda_\mu f_\mu\}_{q_\alpha} = \sum_\mu \lambda_\mu \{E|f_\mu\}_{q_\alpha}. \tag{3.7}$$

The Null Class of the Euler-Lagrange Operator

We now consider to what extent the function $f(q, q_{,i}, \boldsymbol{x})$ is determined by the condition that the Euler-Lagrange Operator $\{E|f\}_q$ vanish. Expanding equation (3.6) in explicit form, we have

$$\{E|f\}_{q_\alpha} = f_{,q_\alpha} - f_{,q_{\alpha,i} q_{\beta,j}} q_{\beta,ji} - f_{,q_{\alpha,i} q_\beta} q_{\beta,i} - \partial_i f_{,q_{\alpha,i}} \tag{3.8}$$

which may be considered as a function of the $N + nN + \frac{1}{2} n N(n+1) + n$ variables $q_\alpha, q_{\alpha,i}, q_{\alpha,ij}, x^i$. Suppose we have two functions $g(q, q_{,i}, \boldsymbol{x})$ and $h(q, q_{,i}, \boldsymbol{x})$ such

that the $2N$ Euler-Lagrange operators $\{E\,|\,g\}_{q_\alpha}$ and $\{E\,|\,h\}_{q_\alpha}$, considered as functions of the $N+nN=\frac{1}{2}nN(n+1)+n$ variables, as above, are equal pair by pair (that is, $\{E\,|\,g\}_{q_\alpha}=\{E\,|\,h\}_{q_\alpha}$), so that $\{E\,|\,g-h\}_{q_\alpha}=0$. Thus, we consider the situation

$$\{E\,|\,k(q_\alpha;\, q_{\alpha,i};\, \boldsymbol{x})\}_{q_\alpha}\equiv 0$$

in $N+nN+\frac{1}{2}nN(n+1)+n$ variables. Such a function k will be said to be an element of the **null class** $\mathcal{N}(E)$ of the Euler-Lagrange operator $\{E\,|\,\mathcal{L}\}$.

For purposes which will be evident below, we shall need to characterize precisely the null-class of the Euler-Lagrange operator $\{E\,|\;\}_{a_\alpha}$.

Theorem 3.2. *Let $a_\alpha(\boldsymbol{x})$ be a collection of functions of class C^2 for all $\boldsymbol{x}$ in a given $\mathcal{D}_n^*$ where the numerosity of the α index set is N. The most general element of the null-class of the Euler-Lagrange operator, $\mathcal{N}(E)$, in ascending powers of $a_{\alpha,i}$, is given by $\sum_{s=0}^{r} O_s(B_{ai}\,a_{a,i})$, where*

(1) $O_s(B_{\alpha i}\,a_{\alpha,i})=B_{\alpha_1 i_1\ldots a_s i_s}a_{\alpha_1,i_1}\cdots a_{\alpha_s,i_s}$

(2) *The B's are of class C^1 and depend on $\boldsymbol{x}$ and a_α but not on $a_{\alpha,i}$.*

(3) $r+1=\max\,(2,\,\min\,(n,N))$

(4) *For $s\geqq 2$, the B's are completely antisymmetric in the Latin and Greek indices separately.*

(5)
$$B_{,a_{\alpha_1}}=\partial_{i_1}B_{\alpha_1 i_1};\qquad s=0 \tag{3.9a}$$

$$B_{\alpha_2 i_1,a_{\alpha_1}}+2\partial_{i_2}B_{\alpha_1 i_1 a_2 i_2}=B_{\alpha_1 i_1,a_{\alpha_2}};\qquad s=1 \tag{3.9b}$$

$$\left.\begin{aligned} B_{\alpha_1 i_1\ldots\alpha_s i_s,a_{\alpha_{s+1}}} &= \partial_{i_{s+1}}B_{\alpha_1 i_1\ldots\alpha_{s+1}i_{s+1}}\\ B_{\alpha_1 i_1\ldots\alpha_s i_s,a_{s+1}} &= -B_{\alpha_1 i_1\ldots\alpha_{s-1}i_{s-1}\alpha_{s+1}i_s,a_{\alpha_s}}\end{aligned}\right\};\qquad 1<s<r \tag{3.9c}$$

$$r\,B_{\alpha_1 i_1\ldots\alpha_{r-1}i_{r-1}\mu\, i_r,a_{\alpha_r}}=B_{\alpha_1 i_1\ldots\alpha_r i_r,a_{\alpha_s}};\qquad s=r. \tag{3.9d}$$

Proof. If $A(a_\alpha;\,a_{\alpha,i};\,\boldsymbol{x})$ is a general element of the null-class of the Euler-Lagrange operator, then by definition

$$a_{\beta,ji}A_{,a_{\alpha,i}a_{\beta,j}}+a_{\beta,i}A_{,a_{\alpha,i}a_\beta}+\partial_i A_{,a_{\alpha,i}}-A_{,a_\alpha}=0$$

must be an identity in $a_{\beta,ij}$, $a_{\beta,i}$, and a_β for all $\boldsymbol{x}$ in a given $\mathcal{D}_n^*$. This identity can be satisfied for all $a_\beta(\boldsymbol{x})$ which are assumed to be of class C^2 only if $a_{\beta,ji}A_{,a_{\alpha,i}a_{\beta,j}}=0$ is an identity in $a_{\beta,ji}$. Due to the assumed continuity conditions on a_β, $a_{\beta,ji}$ is symmetric in i and j, and thus it is sufficient to require

$$A_{,a_{\mu,m}a_{\xi,n}}+A_{,a_{\mu,n}a_{\xi,m}}=0 \tag{3.10}$$

in order to satisfy $a_{\beta,ji}A_{,a_{\alpha,i}a_{\beta,j}}\equiv 0$. Consider the following formal power series in $a_{\alpha,i}$ for A:

$$A=\sum_{s=0}^{\infty}O_s(B_{\alpha i}\,a_{\alpha,i})$$

where

$$O_0=B,\qquad O_k=B_{\alpha_1 i_1\ldots\alpha_k i_k}a_{\alpha_1,i_1}\cdots a_{\alpha_k,i_k}$$

and the B's depend at most on a_α and $\boldsymbol{x}$. From the form of $O_n (n > 1)$ we have

$$B_{\alpha_1 i_1 \ldots \alpha_k i_k \ldots \alpha_l i_l \ldots} = B_{\alpha_1 i_1 \ldots \alpha_l i_l \ldots \alpha_k i_k \ldots} \tag{3.11}$$

for all $k < l \leq n$. It is immediate, upon inspection, that O_0 and O_1 satisfy equation (3.10). By equation (3.11) we have

$$O_{s, a_{\mu, m}} = s\, B_{\alpha_1 i_1 \ldots \alpha_{s-1} i_{s-1} \mu m}\, a_{\alpha_1, i_1} \cdots a_{\alpha_{s-1}, i_{s-1}}$$

and

$$O_{s, a_{\mu, m} a_{\xi, n}} = s(s-1)\, B_{\alpha_1 i_1 \ldots \alpha_{s-2} i_{s-2} \mu m \xi n}\, a_{\alpha_1, i_1} \cdots a_{\alpha_{s-2}, i_{s-2}}.$$

Hence, to satisfy equation (3.10), we require

$$0 = \sum_s s(s-1)\, a_{\alpha_1, i_1} \cdots a_{\alpha_{s-2}, i_{s-2}} \left(B_{\alpha_1 i_1 \ldots \alpha_{s-2} i_{s-2} \mu m \xi n} + B_{\alpha_1 i_1 \ldots \alpha_{s-2} i_{s-2} \mu n \xi m}\right).$$

For this to be true for all $a_{\alpha, i}$, it is required that

$$B_{\alpha_1 i_1 \ldots \alpha_{s-2} i_{s-2} \mu m \xi n} + B_{\alpha_1 i_1 \ldots \alpha_{s-2} i_{s-2} \mu n \xi m} = 0.$$

This result implies, by equation (3.11), that $B_{\alpha_1 i_1 \ldots \alpha_s i_s}$ must be antisymmetric in all Latin and Greek subscripts separately. With N the numerosity of the index set $\mathscr{A}$ and i ranging over an n-dimensional space index, $B_{\alpha_1 i_1 \ldots \alpha_k i_k} = 0$ for all $k \geqq \max[2, \min(n, N)] = r + 1$. We thus obtain the following finite series as the most general form which will satisfy equation (3.10):

$$A = \sum_{s=0}^{r} O_s (B_{\alpha i}\, a_{\alpha_1 i}). \tag{3.12}$$

To satisfy the identical vanishing of the Euler-Lagrange derivative of A it remains to satisfy

$$a_{\xi, m}\, A_{, a_{\mu, m} a_\xi} + \partial_m A_{, a_{\mu, m}} - A_{, a_\mu} = 0. \tag{3.13}$$

Substituting equation (3.12) into equation (3.13) and using the properties of the B's thus far established, we obtain

$$\begin{aligned}\sum_{s=0}^{r} \{ s\, a_{\alpha_1, i_1} \cdots a_{\alpha_{s-1}, i_{s-1}}\, a_{\xi, m} (B_{\alpha_1 i_1 \ldots \alpha_{s-1} i_{s-1} \mu m})_{, \alpha_\xi} &+ \\ + s\, a_{\alpha_1, i_1} \cdots a_{\alpha_{s-1}, i_{s-1}}\, \partial_m (B_{\alpha_1 i_1 \ldots \alpha_{s-1} i_{s-1} \mu m}) &- \\ - a_{\alpha_1, i_1} \cdots a_{\alpha_s, i_s} (B_{\alpha_1, i_1 \ldots \alpha_s i_s})_{, a_\mu} \} &= 0\end{aligned}$$

as an identy to be satisfied in $a_{\alpha, i}$. This identity can be satisfied for all $a_{\alpha, i}$ if and only if the coefficients of each $a_{\alpha, i}$ vanish separately. Thus, upon making the substitution $\mu = \alpha_{s+1}$, we obtain

$$\begin{aligned}s (B_{\alpha_1 i_1 \ldots \alpha_{s-1} i_{s-1} \alpha_{s+1} i_s})_{, a\alpha_s} - (B_{\alpha_1 i_1 \ldots \alpha_s i_s})_{, a\alpha_{s+1}} &+ \\ + (s+1)\, \partial_{i_{s+1}} (B_{\alpha_1 i_1 \ldots \alpha_{s+1} i_{s+1}}) = 0, \qquad s < r,\end{aligned} \tag{3.14}$$

$$r (B_{\alpha_1 i_1 \ldots \alpha_{r-1} i_{r-1} \mu i_r})_{, a\alpha_r} - (B_{\alpha_1 i_1 \ldots \alpha_r i_r})_{, a_\mu} = 0, \qquad s = r. \tag{3.15}$$

If α_s and α_{s+1} are interchanged in the development, we obtain

$$\begin{aligned}s (B_{\alpha_1 i_1 \ldots \alpha_{s-1} i_{s-1} \alpha_s i_s})_{, a\alpha_{s+1}} - (B_{\alpha_1 i_1 \ldots \alpha_{s-1} i_{s-1} \alpha_{s+1} i_s})_{, a\alpha_s} &+ \\ + (s+1)\, \partial_{i_{s+1}} (B_{\alpha_1 i_1 \ldots \alpha_{s-1} i_{s-1} \alpha_{s+1} i_s \alpha_s i_{s+1}}) = 0\end{aligned}$$

for $0<s<r$. Adding this to equation (3.14) and using the antisymmetric properties of the B's with respect to the α's, results in

$$s\{(B_{\alpha_1 i_1\dots\alpha_{s-1} i_{s-1}\alpha_s i_s}),_{a\alpha_{s+1}}+(B_{\alpha_1 i_1\dots\alpha_{s-1} i_{s-1}\alpha_{s+1} i_s}),_{a_s}\}$$
$$=(B_{\alpha_1 i_1\dots\alpha_{s-1} i_{s-1}\alpha_{s+1} i_s}),_{a\alpha_s}+(B_{\alpha_1 i_1\dots\alpha_{s-1} i_{s-1}\alpha_s i_s}),_{a\alpha_{s+1}},\qquad 1<s<r,$$

since for that value of s (*i.e.*, $s=1$) the above equation results in an identity. It is thus required, for $s>1$, that

$$(B_{\alpha_1 i_1\dots\alpha_s i_s}),_{a\alpha_{s+1}}+(B_{\alpha_1 i_1\dots\alpha_{s-1} i_{s-1}\alpha_{s+1} i_s}),_{a\alpha_s}=0. \tag{3.16}$$

Combining these results, the following system of equations is obtained which the B's must satisfy in order for A to be an element of the null-class of the Euler-Lagrange operator:

$$\begin{aligned}
&\partial_{i_1} B_{\alpha_1 i_1}=B_{,\alpha_1},\\
&B_{\alpha_2 i_1, a\alpha_1}+2\partial_{i_2} B_{\alpha_1 i_1\alpha_2 i_2}=B_{\alpha_1 i_1, a\alpha_2},\\
&B_{\alpha_1 i_1\dots\alpha_s i_s, a\alpha_{s+1}}=\partial_{i_{s+1}} B_{\alpha_1 i_1\dots\alpha_s i_s\alpha_{s+1} i_{s+1}}, && 1<s<r,\\
&B_{\alpha_1 i_1\dots\alpha_{s-1} i_{s-1}\alpha_s i_s, a\alpha_{s+1}}=-B_{\alpha_1 i_1\dots\alpha_{s-1} i_{s-1}\alpha_{s+1} i_s, a\alpha_s},\\
&r\,B_{\alpha_1 i_1\dots\alpha_{r-1} i_{r-1}\mu i_r, a\alpha_r}=B_{\alpha_1 i_1\dots\alpha_r i_r, a_\mu},
\end{aligned}$$

which proves the theorem. Q.E.D.

If $r=1$ in Theorem 3.2, we obtain

$$\partial_{i_1} B_{\alpha_1 i_1}=B_{,a\alpha_1},\qquad B_{\alpha_2 i_2, a\alpha_1}=B_{\alpha_1 i_2, a\alpha_2},$$

which are necessary and sufficient conditions for the existence of a variable n-vector function $\mathbf{\Gamma}(a_\alpha,\boldsymbol{x})$ of class C^2 such that

$$B=\partial_i\Gamma^i,\qquad B_{\alpha_1 i_1}=\Gamma^i{}_{,a\alpha}.$$

We shall obtain a particular solution to equations (3.9) for illustrative purposes. Starting with $B=\partial_{i_1}\Gamma_{i_1}$, we obtain by direct operation

$$\begin{aligned}
B_{\alpha_1 i_1}&=\Gamma_{i_1, a\alpha_1}+\partial_{i_2}\Gamma_{\alpha_1 i_1 i_2},\\
B_{\alpha_1 i_1\alpha_2 i_2}&=\Gamma_{\alpha_1 i_1 i_2, a\alpha_2}+\partial_{i_3}\Gamma_{\alpha_1 i_1\alpha_2 i_2 i_3},\\
&\vdots\\
B_{\alpha_1 i_1\dots\alpha_r i_r}&=\Gamma_{\alpha_1 i_1\dots\alpha_{r-1} i_{r-1} i_r, a\alpha_r}
\end{aligned}$$

where

$$\left.\begin{aligned}
&\Gamma_{\alpha_1 i_1\dots\alpha_k i_k i_{k+1}}+\Gamma_{\alpha_1 i_1\dots\alpha_k i_{k+1} i_k}=0\\
&\Gamma_{\alpha_k i_1\dots\alpha_{k-1} i_{k-1} i_k, a\alpha_1}+\Gamma_{\alpha_1 i_1\dots\alpha_{k-1} i_{k-1} i_k, a\alpha_k}=0\\
&\Gamma_{\alpha_1 i_1\dots\alpha_k i_k i_{k+1}, a\alpha_{k+1} a\alpha_{k+2}}=0
\end{aligned}\right\}\qquad k>0,$$

and all Γ's are of class C^2 and antisymmetric in the i's and the α's. Thus if

$$s<r,\quad O_s=(\Gamma_{\alpha_1 i_1\dots\alpha_{s-1} i_{s-1} i_s, a\alpha_s}+\partial_{i_{s+1}}\Gamma_{\alpha_1 i_1\dots\alpha_s i_s i_{s+1}})\,a_{\alpha_1, i_1}\cdots a_{\alpha_s, i_s},$$

and if

$$s=r,\quad O_r=\Gamma_{\alpha_1 i_1\dots\alpha_{r-1} i_{r-1} i_r, a\alpha_r}\,a_{\alpha_1, i_1}\cdots a_{\alpha_r, i_r}.$$

Substituting these results into equation (3.12) of Theorem 3.2 gives

$$A = (\partial_{i_1} \Gamma_1 + \Gamma_{i_1, a\alpha_1} a_{\alpha_1, i_1}) + (\partial_{i_2} \Gamma_{\alpha_1 i_1 i_2} a_{\alpha_1, i_1} + \Gamma_{\alpha_1 i_1 i_2, a\alpha_2} a_{\alpha_1, i_1} a_{\alpha_2, i_2}) + \cdots$$
$$+ (\partial_{i_r} \Gamma_{\alpha_1 i_1 \ldots \alpha_{r-1} i_{r-1} i_r} a_{\alpha_1, i_1} \cdots a_{\alpha_{r-1}, i_{r-1}} + \Gamma_{\alpha_1 i_1 \ldots \alpha_{r-1} i_{r-1} i_r, a\alpha_r} a_{\alpha_1, i_1} \cdots a_{\alpha_r, i_r}).$$

Now

$$(\Gamma_{\alpha_1 i_1 \ldots \alpha_k i_k i_{k+1}} a_{\alpha_1, i_1} a_{\alpha_k, i_k})_{, i_{k+1}}$$
$$= (\partial_{i_{k+1}} \Gamma_{\alpha_1 i_1 \ldots \alpha_k i_k i_{k+1}} + \Gamma_{\alpha_1 i_1 \ldots \alpha_k i_k i_{k+1}, a_{k+1}}) a_{\alpha_1, i_1} \cdots a_{a_{k+1}, i_{k+1}} +$$
$$+ k \, \Gamma_{\alpha_1 i_1 \ldots \alpha_k i_k i_{k+1}} a_{\alpha_1, i_1} \cdots a_{\alpha_k, i_k i_{k+1}}$$

of which the second term on the right vanishes due to the antisymmetry of the Γ's and the symmetry of $a_{\alpha_k, i_k i_{k+1}}$ in i_k and i_{k+1}. Hence, upon rearrangement, we obtain

$$A = (\Gamma_{i_r} + \Gamma_{\alpha_1 i_1 i_r} a_{\alpha_1, i_1} + \cdots + \Gamma_{\alpha_1 i_1 \ldots \alpha_{r-1} i_{r-1} i} \, a_{\alpha_1, i_1} \cdots a_{\alpha_{r-1}, i_{r-1}})_{, i_r}. \tag{3.17}$$

The usual statement[10] made concerning the null-class of the Euler-Lagrange operator is that a divergence is variationally deletable. The above particular solution bears out this statement only under the very special conditions placed upon the Γ's. In other words an arbitrary divergence is not necessarily variationally deletable. The special form of the A above can be interpreted as requiring that the divergence shall not explicitly contain second derivatives.

It is thus evident that if two functions $g(\mathrm{q}; \mathrm{q}_{,i}; \boldsymbol{x})$ and $h(\mathrm{q}; \mathrm{q}_{,i}; \boldsymbol{x})$ are such that $\{E|g\}_{\mathrm{q}} = \{E|h\}_{\mathrm{q}}$, then there exists a function $U(\mathrm{q}; \mathrm{q}_{,i}; \boldsymbol{x})$, which is an element of the null-set of the Euler-Lagrange operator, such that $h = g + U$.

Correspondence between Lagrangean and Hamiltonian Systems

Equations (3.3) are N second order partial differential equations for the determination of the N functions $\mathrm{q}_\alpha(\boldsymbol{x})$ which are such that the functional

$$\mathscr{I} = \int_{\mathscr{D}_n^*} f(\mathrm{q}_\alpha; \mathrm{q}_{\alpha, i}; \boldsymbol{x}) \, dv_n$$

has an extremal value. We now wish to decompose these equations by variational methods into a particular set of $N(n+1)$ partial differential equations of the first order which are such that all non-linearities are of an algebraic nature only. We commence by defining nN new variables $p_{\alpha i}$ by the equations

$$p_{\alpha i} \stackrel{\text{def}}{=} f_{, \mathrm{q}_{\alpha, i}}. \tag{3.18}$$

Under this substitution, we have from equation (3.2)

$$\{\varepsilon|f\}_{\mathrm{q}_\alpha} = f_{, \mathrm{q}_\alpha} - p_{\alpha i, i}. \tag{3.19}$$

It is noted that equation (3.18) yields a well defined transformation $\mathrm{q}_{\alpha, i} \to p_{\alpha i}$ if the condition $\det|f_{, \mathrm{q}_{\alpha, i} \mathrm{q}_{\beta, j}}| \neq 0$ is satisfied.

Consider the function H defined in terms of the $N + 2nN + n$ variables $\mathrm{q}_\alpha; \mathrm{q}_{\alpha, i}; p_{\alpha i}, \boldsymbol{x}$ by

$$H \stackrel{\text{def}}{=} p_{\alpha i} \mathrm{q}_{\alpha, i} - f(\mathrm{q}_\alpha; \mathrm{q}_{\alpha, i}; \boldsymbol{x}). \tag{3.20}$$

[10] LANCZOS, C.: Rev. Mod. Phys. **29**, No. 3, 339 (July 1957).

Computing the differential of H, we have

$$\begin{aligned} dH = \mathrm{q}_{\alpha,i}\, d p_{\alpha i} + p_{\alpha i}\, d\mathrm{q}_{\alpha,i} \\ - f_{,\mathrm{q}_\alpha}\, d\mathrm{q}_\alpha - f_{,\mathrm{q}_{\alpha,i}}\, d\mathrm{q}_{\alpha,i} \\ - \partial_i f\, d x_i, \end{aligned}$$

which by equation (3.18) gives

$$dH = \mathrm{q}_{\alpha,i}\, d p_{\alpha i} - f_{,\mathrm{q}_\alpha}\, d\mathrm{q}_\alpha - \partial_i f\, d x^i,$$

so that H may be considered as a function of $p_{\alpha i}$; q_α; $\boldsymbol{x}$. (It is noted that H is of class C^2 in its arguments, since f is, as seen by equation (3.20). From the differential of H we have

$$\begin{aligned} \mathrm{q}_{\alpha,i} &= H_{,p_{\alpha i}}, \\ -f_{,\mathrm{q}_\alpha} &= H_{,\mathrm{q}_\alpha}, \\ -\partial_i f &= \partial_i H, \end{aligned} \tag{3.21}$$

and hence equation (3.19) becomes

$$\{E \mid f\}_{\mathrm{q}_\alpha} = -H_{,\mathrm{q}_\alpha} - p_{\alpha i,i}. \tag{3.22}$$

Upon combining the above results, we conclude that

$$\mathrm{q}_{\alpha,i} = H_{,p_{\alpha i}} \qquad p_{\alpha i,i} = -H_{,\mathrm{q}_\alpha} \tag{3.23}$$

which states that (p, q) form a Hamiltonian system base H, possessing the desired property of having only algebraic nonlinearities.

If the matrix $(f_{,\mathrm{q}_{\alpha,i}\mathrm{q}_{\beta,j}}) = (p_{\alpha i},\, \mathrm{q}_{\beta,j})$ is nonsingular, we may solve, at least locally, for $\mathrm{q}_{\beta,j}$ as a function of $p_{\alpha i}$; q_α; $\boldsymbol{x}$ by equation (3.18) and thereby obtain H as a function of $(p, \mathrm{q}, \boldsymbol{x})$ from equation (3.20). It is evident that if $(f_{,\mathrm{q}_{\alpha,i}\mathrm{q}_{\beta,j}})$ is nonsingular, then $(H_{,p_{\alpha i}p_{\beta j}})$ is nonsingular since

$$(H_{,p_{\alpha i}p_{\beta j}}) = (\mathrm{q}_{\alpha,i,p_{\beta j}}) = (p_{\alpha i,\mathrm{q}_{\beta,j}})^{-1} = (f_{,\mathrm{q}_{\alpha,i}\mathrm{q}_{\beta,j}})^{-1}. \tag{3.24}$$

Consider now the function f defined by

$$f(\mathrm{q}_\alpha;\, \mathrm{q}_{\alpha,i};\, p_{\alpha i};\, p_{\alpha i,j};\, \boldsymbol{x}) = p_{\alpha i}\, \mathrm{q}_{\alpha,i} - H(p, \mathrm{q}, \boldsymbol{x}).$$

Then

$$df = p_{\alpha i}\, d\mathrm{q}_{\alpha,i} + \mathrm{q}_{\alpha,i}\, d p_{\alpha i} - H_{,p_{\alpha i}}\, d p_{\alpha i} - H_{,\mathrm{q}_\alpha}\, d\mathrm{q}_\alpha - \partial_i H\, d x^i.$$

If we consider $p_{\alpha i}$ as a function of q_α; $\mathrm{q}_{\alpha,i}$; $\boldsymbol{x}$ defined by

$$\mathrm{q}_{\alpha,i} = H_{,p_{\alpha i}},$$

which is possible if $\det |H_{,p_{\alpha i}p_{\beta j}}| \neq 0$, then

$$df = p_{\alpha i}\, d\mathrm{q}_{\alpha,i} - H_{,\mathrm{q}_\alpha}\, d\mathrm{q}_\alpha - \partial_i H\, d x^i,$$

and hence

$$f_{,\mathrm{q}_{\alpha,i}} = p_{\alpha i}; \qquad f_{,\mathrm{q}_\alpha} = -H_{,\mathrm{q}_\alpha}; \qquad \partial_i f = -\partial_i H.$$

Comparison of the above result with equation (3.22) shows that we recover the Euler-Lagrange equations. We have thus proved

Theorem 3.3. *The Euler-Lagrange equations and the Hamiltonian equations are equivalent provided the matrix* $(f_{,q_{\alpha,i}q_{\beta,j}})$ *is nonsingular, in which case*

$$(f_{,q_{\alpha,i}q_{\beta,j}}) = (H_{,p_{\alpha i}p_{\beta j}})^{-1}.$$

Canonical Maps and Variational Considerations

Consider the function h of the $N+2nN+n^2N+n$ variables $q_\alpha; q_{\alpha,i}; p_{\alpha i}; p_{\alpha i,j}; \boldsymbol{x}$ as defined by

$$h = p_{\alpha i} q_{\alpha,i} - H(p, q, \boldsymbol{x}); \tag{3.25}$$

then

$$\{E|h\}_{q_\alpha} = -H_{,q_\alpha} - p_{\alpha i,i}, \tag{3.26}$$

$$\{E|h\}_{p_{\alpha i}} = q_{\alpha,i} - H_{,p_{\alpha i}}. \tag{3.27}$$

Hence, for such an h, the Euler-Lagrange operator results in the corresponding Hamiltonian system for the corresponding f.

Consider a transformation of the collection $\mathscr{T}$ on the variables (p, q). We reformulate the definition of a canonical map in terms of Euler-Lagrange operators.

A transformation T of the collection $\mathscr{T}$ is said to be **canonical** if and only if there exists for every $H\in\mathfrak{H}$ a $K\in\mathfrak{H}$ such that

$$\begin{aligned}\mu\{E|h\}_{Q_\alpha} &= -K_{,Q_\alpha} - P_{\alpha i,i},\\ \mu\{E|h\}_{P_{\alpha i}} &= Q_{\alpha,i} - K_{,P_{\alpha i}},\end{aligned} \tag{3.28}$$

where h is defined by equation (3.25) and μ is a nonzero constant.

The equivalence of this definition and the previous definition of a canonical map is seen as follows. By Theorem 3.1, the Euler-Lagrange operator is invariant under all T of the collection $\mathscr{T}$. Thus, by equations (3.26) and (3.27), if (p, q) is to form a Hamiltonian system,

$$\{E|h\}_{q_\alpha} = 0, \quad \{E|h\}_{p_{\alpha i}} = 0,$$

then

$$\{E|h\}_{Q_\alpha} = 0, \quad \{E|h\}_{P_{\alpha i}} = 0$$

so that (P, Q) form a Hamiltonian system by equations (3.28).

We are now in possession of the results required to obtain the alternate characterization of canonical maps.

Theorem 3.4. *A map T of the collection $\mathscr{T}$ is a canonical map if and only if there exists a nonzero finite constant μ and a* K *for every* $H\in\mathfrak{H}\ominus\mathfrak{H}_0$ *such that*

$$\mu\,(p_{\alpha i} q_{\alpha,i} - H(p, q, \boldsymbol{x}))$$

differs from

$$P_{\alpha i} Q_{\alpha,i} - K(P, Q, \boldsymbol{x})$$

by at most an element of the null set of the Euler-Lagrange operators

$$\{E|\ \}_{P_{\alpha i}} \text{ and } \{E|\ \}_{Q_\alpha} \text{ for all } \boldsymbol{x} \text{ in } \mathscr{D}_n^*.$$

Proof. Consider the function $\varphi(P_{\alpha i}; P_{\alpha i,j}; Q_\alpha; Q_{\alpha i}; \boldsymbol{x})$ defined by

$$\varphi = P_{\alpha i} Q_{\alpha,i} - K(P, Q, \boldsymbol{x}).$$

Application of the Euler-Lagrange operators $\{E|\ \}_{P_{\alpha i}}$ and $\{E|\ \}_{Q_\alpha}$ results in

$$\{E|\varphi\}_{Q_\alpha} = -\mathrm{K}_{,Q_\alpha} - P_{\alpha i,i}$$
$$\{E|\varphi\}_{P_{\alpha i}} = Q_{\alpha,i} - \mathrm{K}_{,P_{\alpha i}}.$$

Comparing these results with the defining equations of a canonical map of this section (that is, equation (3.28)), we have

$$\mu\{E|h\}_{Q_\alpha} = \{E|\varphi\}_{Q_\alpha}$$
$$\mu\{E|h\}_{P_{\alpha i}} = \{E|\varphi\}_{P_{\alpha i}}$$

in order for T to be canonical. This condition can be satisfied only if $\mu h - \varphi$ is equal to an element of the null class of the Euler-Lagrange operators $\{E|\ \}_{Q_\alpha}$ and $\{E|\ \}_{P_{\alpha i}}$ as seen above (*cf.* p. 124). Q.E.D.

Equivalence of the Two Characterizations of Canonical Maps

We now proceed to show that this alternate characterization leads to the same conditions on a map T in order for it to be canonical as obtained in Section II, namely that the associated Jacobian matrix of T must be a generalized symplectic matrix.

From Theorem 3.4 we see that a necessary and sufficient condition for a map $T: p_{\alpha i} = p_{\alpha i}(P, Q, \boldsymbol{x})$, $\mathrm{q}_\alpha = \mathrm{q}_\alpha(P, Q, \boldsymbol{x})$ to be canonical is that

$$\mu\left(p_{\alpha i}\,\mathrm{q}_{\alpha,i} - H(p, \mathrm{q}, \boldsymbol{x})\right) - P_{\beta j}\,Q_{\beta,j} + \mathrm{K}(P, Q, \boldsymbol{x})$$

differs from zero by at most an element of the null-class of $\{E|\ \}_{PQ}$, where (p, q) are considered functions of the (P, Q). We have

$$\left(\mu\left(p_{\alpha i}\,\mathrm{q}_{\alpha,i} - H\right) - P_{\beta j}\,Q_{\beta,j} + \mathrm{K}\right) \in \mathcal{N}(E).$$

From equation (2.2), the derivatives of q_α may be calculated considering q_α as a function of $P, Q, \boldsymbol{x}$; the result is

$$\mathrm{q}_{\alpha,i} = \partial_i\,\mathrm{q}_\alpha + \mathrm{q}_{\alpha,P_{\lambda k}} P_{\lambda k,i} + \mathrm{q}_{\alpha,Q_\lambda} Q_{\lambda,i}.$$

From Theorem 3.2 we have

$$\mathcal{N}(E) = B + B_{\lambda i}\,Q_{\lambda,i} + B_{\underset{\smile}{\lambda k}\,j}\,P_{\lambda k,j} + B_{\lambda i\beta j}\,Q_{\lambda,i}\,Q_{\beta,j} + \cdots$$

where the index set of the a_α of Theorem 3.2 is such that the a_α are in a one-to-one correspondence with the collection of functions Q_α, $P_{\alpha i}$ and where the B's satisfy equations (3.9). Hence,

$$\begin{aligned}\mu\,p_{\alpha i}\left(\partial_i\,\mathrm{q}_\alpha + \mathrm{q}_{\alpha,P_{\lambda k}} P_{\lambda k,i} + \mathrm{q}_{\alpha,Q_\lambda} Q_{\lambda,i}\right) - \mu H - P_{\beta j}\,Q_{\beta,j} + \mathrm{K}\\ = B + B_{\alpha i}\,Q_{\alpha,i} + B_{\underset{\smile}{\alpha j}\,i}\,P_{\alpha j,i} + B_{\alpha i\beta j}\,Q_{\alpha,i}\,Q_{\beta,j} +\\ + B_{\underset{\smile}{\alpha i}\,j\,\underset{\smile}{\beta k}\,l}\,P_{\alpha i,j}\,P_{\beta k,l} + 2B_{\underset{\smile}{\alpha i}\,j\beta k}\,P_{\alpha i,j}\,Q_{\beta,k} + \cdots\end{aligned}$$

($\underset{\smile}{\alpha i}$ means that the range of the B's in Theorem 3.2 is such as to include α and i in one index set of the Greek suffices). If the above equation is to be satisfied for P, Q which are independent of H, the coefficients of the derivatives of P, Q

must vanish separately, giving

$$\mu\, p_{\alpha i}\, \partial_i\, q_\alpha - \mu H + K = B, \tag{3.29}$$

$$\mu\, p_{\alpha i}\, q_{\alpha, Q_\lambda} - P_{\lambda i} = B_{\lambda i}, \tag{3.30}$$

$$\mu\, p_{\alpha i}\, q_{\alpha, P_{\lambda k}} = B_{\underset{\smile}{\lambda k} i}, \tag{3.31}$$

and all other B's must vanish for all values of their indices. Thus equations (3.9) reduce to

$$B_{,a_\alpha} = \partial_i B_{\alpha i},$$

$$B_{\alpha i, a_\beta} = B_{\beta i, a_\alpha}$$

which implies that there exists an n-vector function $\mathbf{\Gamma}$ of class C^2 such that

$$B = \partial_i \Gamma_i, \qquad B_{\alpha i} = \Gamma_{i, a_\alpha},$$

and hence, under the identification of the a_α with (P, Q), we obtain

$$\begin{aligned} \partial_i \Gamma_i &= \mu\, p_{\alpha i}\, \partial_i\, q_\alpha - \mu H + K, \\ \Gamma_{i, Q_\lambda} &= \mu\, p_{\alpha i}\, q_{\alpha, Q_\lambda} - P_{\lambda i}, \\ \Gamma_{i, P_{\lambda k}} &= \mu\, p_{\alpha i}\, q_{\alpha, P_{\lambda k}}. \end{aligned} \tag{3.32}$$

For these equations to possess a solution, it is both necessary and sufficient that

$$(\partial_i \Gamma_i)_{, Q_\alpha} = \partial_i (\Gamma_{i, Q_\lambda}), \tag{3.33}$$

$$(\partial_i \Gamma_i)_{, P_{\lambda k}} = \partial_i (\Gamma_{i, P_{\lambda k}}), \tag{3.34}$$

$$(\Gamma_{i, Q_\lambda})_{, Q_\beta} = (\Gamma_{i, Q_\beta})_{, Q_\lambda}, \tag{3.35}$$

$$(\Gamma_{i, Q_\beta})_{, P_{\lambda k}} = (\Gamma_{i, P_{\lambda k}})_{, Q_\beta}, \tag{3.36}$$

$$(\Gamma_{i, P_{\lambda k}})_{, P_{\beta l}} = (\Gamma_{i, P_{\beta l}})_{, P_{\lambda k}}, \tag{3.37}$$

where the Γ's stand for the corresponding right sides of equations (3.32).

Equation (3.33) gives, upon using equation (3.32),

$$\mu\, q_{\alpha, Q_\lambda}\, \partial_i\, p_{\alpha i} + \mu\, p_{\alpha i}\, \partial_i\, q_{\alpha, Q_\lambda} = \mu\, p_{\alpha i, Q_\lambda}\, \partial_i\, q_\alpha - \mu H_{, Q_\lambda} + K_{, Q_\lambda}. \tag{3.38}$$

Now

$$H_{, Q_\lambda} = H_{, p_{\alpha i}}\, p_{\alpha i, Q_\lambda} + H_{, q_\alpha}\, q_{\alpha, Q_\lambda}$$

which by the use of $H_{, p_{\alpha i}} = q_{\alpha, i}$, $H_{, q_\alpha} = -p_{\alpha i, i}$ becomes

$$H_{, Q_\lambda} = q_{\alpha, i}\, p_{\alpha i, Q_\lambda} - p_{\alpha i, i}\, q_{\alpha, Q_\lambda}.$$

However

$$p_{\alpha i, i} = p_{\alpha i, Q_\beta}\, Q_{\beta, i} + p_{\alpha i, P_{\beta j}}\, P_{\beta j, i} + \partial_i\, p_{\alpha i},$$

$$q_{\alpha, i} = q_{\alpha, Q_\beta}\, Q_{\beta, i} + q_{\alpha, P_{\beta j}}\, P_{\beta j, i} + \partial_i\, q_\alpha,$$

so that

$$H_{,Q_\lambda} = p_{\alpha i,Q_\lambda}(\mathrm{q}_{\alpha,Q_\beta} Q_{\beta,i} + \mathrm{q}_{\alpha,P_{\beta j}} P_{\beta j,i} + \partial_i \mathrm{q}_\alpha) -$$
$$- \mathrm{q}_{\alpha,Q_\lambda}(p_{\alpha i,Q_\beta} Q_{\beta,i} + p_{\alpha i,P_{\beta j}} P_{\beta j,i} + \partial_i p_{\alpha i}).$$

Substituting these results into equation (3.38) results in, after cancellation of like terms,

$$\mu\, \mathrm{q}_{\alpha,Q_\lambda}(p_{\alpha i,Q_\beta} Q_{\beta,i} + p_{\alpha i,P_{\beta j}} P_{\beta j,i}) - \mu\, p_{\alpha i,Q_\lambda}(\mathrm{q}_{\alpha,Q_\beta} Q_{\beta,i} + \mathrm{q}_{\alpha,P_{\beta j}} P_{\beta j,i}) + \mathrm{K}_{,Q_\lambda} = 0.$$

To simplify the writing of this and subsequent equations, we introduce the **Lagrange-bracket vectrix components** defined by

$$\{i|u, v; p\,\mathrm{q}\} = p_{\alpha i,u}\, \mathrm{q}_{\alpha,v} - p_{\alpha i,v}\, \mathrm{q}_{\alpha,u}.$$

This operator will be dicussed in detail in Section IV.

With the use of the above-defined vectrix operator, the foregoing equation may be written as

$$\mu\, P_{\beta j,i}\{i|P_{\beta j}, Q_\lambda; p\,\mathrm{q}\} + \mu\, Q_{\beta,i}\{i|Q_\beta, Q_\lambda; p\,\mathrm{q}\} + \mathrm{K}_{,Q_\lambda} = 0. \tag{3.39}$$

In a similar manner equation (3.34) results in

$$\mu\, P_{\lambda k,i}\{i|P_{\lambda k}, P_{\beta l}; p\,\mathrm{q}\} + \mu\, Q_{\beta,i}\{i|Q_\beta, P_{\lambda k}; p\,\mathrm{q}\} + \mathrm{K}_{,P_{\lambda k}} = 0. \tag{3.40}$$

Proceeding in the same manner, equations (3.35), (3.36), and (3.37) result in

$$\begin{aligned}
&\mathrm{q}_{\alpha,P_{\beta k}}\, p_{\alpha i,P_{\beta j}} - \mathrm{q}_{\alpha,P_{\beta j}}\, p_{\alpha i,P_{\lambda k}} = 0,\\
&\mathrm{q}_{\alpha,P_{\lambda k}}\, p_{\alpha i,Q_\beta} - p_{\alpha i,P_{\lambda k}}\, \mathrm{q}_{\alpha,Q_\beta} = -\frac{1}{\mu}\, \delta_{\lambda\beta}\, \delta_{ik},\\
&\mathrm{q}_{\alpha,Q_\lambda}\, p_{\alpha i,Q_\beta} - \mathrm{q}_{\alpha,Q_\beta}\, p_{\alpha i,Q_\lambda} = 0
\end{aligned}$$

which, when expressed in terms of Lagrange-bracket-operators, give

$$\begin{aligned}
&\{i|P_{\lambda k}, P_{\beta j}; p\,\mathrm{q}\} = 0,\\
&\{i|P_{\lambda k}, Q_\beta; p\,\mathrm{q}\} = \frac{1}{\mu}\, \delta_{\lambda\beta}\, \delta_{ik},\\
&\{i|Q_\beta, Q_\lambda; p\,\mathrm{q}\} = 0.
\end{aligned} \tag{3.41}$$

Noting the relation

$$\{i|u, v; p\,\mathrm{q}\} = -\{i|v, u; p\,\mathrm{q}\}$$

which follows immediately from the definition of $\{i|u, v; p\mathrm{q}\}$, we have, from equations (3.39) and (3.40),

$$\begin{aligned}
P_{\beta j,i}\, \delta_{\beta\lambda}\, \delta_{ij} + \mathrm{K}_{,Q_\lambda} &= 0,\\
-Q_{\beta,i}\, \delta_{ik}\, \delta_{\beta\lambda} + \mathrm{K}_{,P_{\lambda k}} &= 0.
\end{aligned}$$

These are just the equations required for the (P, Q) variables to be canonical. Since equations (3.41) result in the Hamiltonian equations for (P, Q) for any $H \in \mathfrak{H} \ominus \mathfrak{H}_0$, under the assumption that (p, q) form a Hamiltonian system, we have proved

Theorem 3.5. *A map T of the collection $\mathscr{T}$ is canonical if and only if*

$$\{i|P_{\lambda k}, P_{\beta i}; p\,q\} = 0, \tag{3.42}$$

$$\{i|P_{\lambda k}, Q_\beta; p\,q\} = \frac{1}{\mu}\,\delta_{\lambda\beta}\,\delta_{ik}, \tag{3.43}$$

$$\{i|Q_\beta, Q_\lambda; p\,q\} = 0, \tag{3.44}$$

where μ is a nonzero constant and

$$\{i|u, v; p\,q\} = p_{\alpha i,u}\, q_{\alpha,v} - p_{\alpha i,v}\, q_{\alpha,u}. \tag{3.45}$$

We now show that equations (3.42) through (3.45) are equivalent to the statement that the associated Jacobian matrix of the map T is a generalized symplectic matrix.

Theorem 3.6. *Equations* (3.42) *to* (3.44) *imply*

$$\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} = \frac{1}{\mu}\vec{\boldsymbol{I}}, \tag{3.46}$$

and conversely.

Proof. Expanding the left side of equation (3.46), using equations (2.15) through (2.19) and (2.3), gives

$$\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} = \begin{pmatrix} \boldsymbol{A}'\vec{\boldsymbol{E}}_2\boldsymbol{C} + \boldsymbol{C}'\vec{\boldsymbol{E}}_3\boldsymbol{A} & \boldsymbol{A}'\vec{\boldsymbol{E}}_2\boldsymbol{D} + \boldsymbol{C}'\vec{\boldsymbol{E}}_3\boldsymbol{B} \\ \boldsymbol{B}'\vec{\boldsymbol{E}}_2\boldsymbol{C} + \boldsymbol{D}'\vec{\boldsymbol{E}}_3\boldsymbol{A} & \boldsymbol{B}'\vec{\boldsymbol{E}}_2\boldsymbol{D} + \boldsymbol{D}'\vec{\boldsymbol{E}}_3\boldsymbol{B} \end{pmatrix}.$$

Now

$$\begin{aligned} &\boldsymbol{A}'\vec{\boldsymbol{E}}_2\boldsymbol{C} + \boldsymbol{C}'\vec{\boldsymbol{E}}_3\boldsymbol{A} \\ &\qquad = (q_{\alpha,Q_\beta})'(p_{\alpha 1,Q_\beta}\vec{e}_1 + \cdots + p_{\alpha n,Q_\beta}\vec{e}_n) - \\ &\qquad\quad - [(p_{\alpha 1,Q_\beta})', \ldots, (p_{\alpha n,Q_\beta})'] \begin{Bmatrix} (q_{\alpha,Q_\beta})\vec{e}_1 \\ \vdots \\ (q_{\alpha,Q_\beta})\vec{e}_n \end{Bmatrix} \\ &\qquad = (\{1|Q_\alpha, Q_\beta; p\,q\}\vec{e}_1 + \{2|Q_\alpha, Q_\beta; p\,q\}\vec{e}_2 + \\ &\qquad\quad + \{3|Q_\alpha, Q_\beta; p\,q\}\vec{e}_3 + \cdots + \{n|Q_\alpha, Q_\beta|p\,q\}\vec{e}_n), \end{aligned} \tag{3.47}$$

and

$$\begin{aligned} &\boldsymbol{A}'\vec{\boldsymbol{E}}_2\boldsymbol{D} + \boldsymbol{C}'\vec{\boldsymbol{E}}_3\boldsymbol{B} \\ &\qquad = [(\{1|P_{\alpha 1}, Q_\beta; p\,q\}\vec{e}_1 + \cdots + \{n|P_{\alpha 1}, Q_\beta; p\,q\}\vec{e}_n), \\ &\qquad\quad (\{1|P_{\alpha 2}, Q_\beta; p\,q\}\vec{e}_1 + \cdots + \{n|P_{\alpha 2}, Q_\beta; p\,q\}\vec{e}_n), \\ &\qquad\quad \vdots \\ &\qquad\quad (\{1|P_{\alpha n}, Q_\beta; p\,q\}\vec{e}_1 + \cdots + \{n|P_{\alpha n}, Q_\beta; p\,q\}\vec{e}_n)]. \end{aligned} \tag{3.48}$$

Similarly, since $\boldsymbol{B}'\vec{\boldsymbol{E}}_2\boldsymbol{C} + \boldsymbol{D}'\vec{\boldsymbol{E}}_3\boldsymbol{A} = -(\boldsymbol{A}'\vec{\boldsymbol{E}}_2\boldsymbol{D} + \boldsymbol{C}'\vec{\boldsymbol{E}}_3\boldsymbol{B})'$,

$$\boldsymbol{B}'\vec{\boldsymbol{E}}_2\boldsymbol{C} + \boldsymbol{D}'\boldsymbol{E}_3\boldsymbol{A} = -\begin{Bmatrix} (\{1|P_{\alpha 1}, Q_\beta; p\,q\}\vec{e}_1 + \cdots + \{n|P_{\alpha 1}, Q_\beta; p\,q\}\vec{e}_n)' \\ \vdots \\ (\{1|P_{\alpha n}, Q_\beta; p\,q\}\vec{e}_1 + \cdots + \{n|P_{\alpha n}, Q_\beta; p\,q\}\vec{e}_n)' \end{Bmatrix}. \tag{3.49}$$

Finally

$$\boldsymbol{B}'\vec{\boldsymbol{E}}_2\boldsymbol{D}+\boldsymbol{D}'\vec{\boldsymbol{E}}_3\boldsymbol{B}=\begin{pmatrix}\vec{e}_i(\{i|P_{\alpha 1},P_{\beta 1};p\,\mathrm{q}) & \dots & \vec{e}_i(\{i|P_{\alpha n},P_{\beta 1};p\,\mathrm{q})\\ \vdots & \ddots & \vdots\\ \vec{e}_i(\{i|P_{\alpha 1},P_{\beta n};p\,\mathrm{q}) & \dots & \vec{e}_i(\{i|P_{\alpha n},P_{\beta n};p\,\mathrm{q})\end{pmatrix}. \tag{3.50}$$

The theorem thus follows directly the substitution of equations (3.42) through (3.45), which are necessary and sufficient conditions for T to be canonical with multiplier μ, in equations (3.47) through (3.50) above. Q.E.D.

A very interesting result can be obtained if we compare the results of Theorem 3.5 with the results of the consideration of the remainder function in Section II. From equation (3.29) we have

$$\mathrm{K}=\mu H+B-\mu\, p_{\alpha i}\,\partial_i \mathrm{q}_\alpha.$$

Comparing this with equation (2.49), that is,

$$\mathrm{K}=\mu H-U,$$

we have

$$U=\mu\, p_{\alpha i}\,\partial_i\mathrm{q}_\alpha-B.$$

We have seen, however, that

$$B=\partial_i\Gamma_i(P,Q,\boldsymbol{x}),$$

and hence

$$U=\mu\, p_{\alpha i}\,\partial_i\mathrm{q}_\alpha-\partial_i\Gamma_i(P,Q,\boldsymbol{x}). \tag{3.51}$$

The development given in this section is similar in form and organization to the development given by SIEGEL[11] for the case $n=1$. The only literature concerning canonical maps on Hamiltonian systems for $n>1$ is given by LANCZOS[12]; he requires that the form $p_{\alpha i}\,\mathrm{q}_{\alpha,i}$ change by at most a divergence.

Alpha Systems and the Calculus of Variations

The results obtained to this point have been for systems of equations of the form (2.1). The vectrix $\vec{\boldsymbol{I}}$ which has been fundamental in the development thus far was specifically tailored to fit systems of equations of the form (2.1). We shall show that a majority of the results carry over to cases where $\vec{\boldsymbol{I}}$ is replaced by a vectrix of a different form which is suggested by variational considerations.

Let β be an index which ranges over an index set $\mathscr{B}$ of finite numerosity B, and consider the functions $\Psi_\beta(\boldsymbol{x})$ arranged in a column matrix $\{\Psi\}$, where Ψ_β are of class C^2 in a $\mathscr{D}_n^*$ under consideration, and are such that they extremize the integral

$$\int_{\mathscr{D}_n^*}\mathscr{L}(\Psi_\beta;\Psi_{\beta,i};\boldsymbol{x})\,dv_n, \tag{3.52}$$

where $\mathscr{L}$ is given by

$$\mathscr{L}=\{\Psi\}'\vec{\boldsymbol{A}}\bullet\vec{\nabla}\{\Psi\}-H(\Psi_\beta,\boldsymbol{x}), \tag{3.53}$$

$\vec{\boldsymbol{A}}$ is a constant $B\times B$ vectrix, $\vec{\nabla}$ is the vectrix operator introduced in Section II, and H is of class C^2 in its arguments. Applying the Euler-Lagrange operator

[11] SIEGEL, C. L.: Vorlesungen über Himmelsmechanik, p. 1—11.
[12] LANCZOS, C.: Rev. Mod. Phys. **29**, No. 3, 61 (July 1957).

to $\mathscr{L}$ and setting it to zero, which is a necessary and sufficient condition for the extremization of $\mathscr{I}$, (3.53) gives

$$\vec{\nabla}\{\Psi\}' \bullet (\vec{A} - \vec{A}') + H_{,\{\Psi\}} = [0]. \tag{3.54}$$

The form of $\mathscr{L}$ given by equation (3.53) contains a superfluous symmetric term which we now proceed to eliminate. Separating the vectrix $\vec{A}$ into symmetric and antisymmetric parts, we obtain

$$\{\Psi\}' \vec{A} \bullet \vec{\nabla}\{\Psi\} = \tfrac{1}{2}\{\Psi\}' (\vec{A} + \vec{A}') \bullet \vec{\nabla}\{\Psi\} + \tfrac{1}{2}\{\Psi\}' (\vec{A} - \vec{A}') \bullet \vec{\nabla}\{\Psi\}.$$

Now

$$\{E|\{\Psi\}' (\vec{A} + \vec{A}') \bullet \vec{\nabla}\{\Psi\}\}_{\{\Psi\}} = \vec{\nabla}\{\Psi\}' \bullet (\vec{A} + \vec{A}' - \vec{A} - \vec{A}') = [0],$$

so that the Lagrangean given by equation (3.53) may be replaced by

$$\mathscr{L} = \tfrac{1}{2}\{\Psi\}' (\vec{A} - \vec{A}') \bullet \vec{\nabla}\{\Psi\} - H(\Psi, \boldsymbol{x}) \tag{3.55}$$

since they both result in the same system of equations for the determination of $\{\Psi\}$, namely equations (3.54).

Set

$$\vec{\alpha} = \vec{A} - \vec{A}'; \tag{3.56}$$

then using equation (3.56), equations (3.55) and (3.54) become

$$\vec{\nabla}\{\Psi\}' \bullet \vec{\alpha} + H_{,\{\Psi\}} = [0], \tag{3.57}$$

$$\mathscr{L} = \tfrac{1}{2}\{\Psi\}' \vec{\alpha} \bullet \vec{\nabla}\{\Psi\} - H(\Psi, \boldsymbol{x}). \tag{3.58}$$

Assuming $\vec{\alpha} \bullet \vec{\alpha}$ is nonsingular, it is evident from equation (3.57) that $\{\Psi\}$ forms an alpha system with structure vectrix $\vec{\alpha}$ (*cf.* Section II, equation (2.13)). We thus have the following

Theorem 3.7. *Every system of equations resulting from the extremization of an integral of the form* $\int_{\mathscr{D}_n^*} \mathscr{L} dv_n$ *where* $\mathscr{L}$ *is given by*

$$\mathscr{L} = \{\Psi\}' \vec{A} \bullet \vec{\nabla}\{\Psi\} - H(\Psi, \boldsymbol{x}), \tag{3.59}$$

$\vec{A}$ *is a constant vectrix, and* $\det|(\vec{A} - \vec{A}') \bullet (\vec{A} - \vec{A}')| \neq 0$, *forms an alpha system with structure vectrix*

$$\vec{\alpha} = \vec{A} - \vec{A}', \tag{3.60}$$

and $\mathscr{L}$ *may be replaced by*

$$'\mathscr{L} = \tfrac{1}{2}\{\Psi\}' \vec{\alpha} \bullet \vec{\nabla}\{\Psi\} - H(\Psi, \boldsymbol{x}) \tag{3.61}$$

with no loss of generality.

Let us examine the Hamiltonian system from the standpoint of its being a special case of an alpha system. We have seen that the Lagrangean for a Hamiltonian system was given by

$$\mathscr{L} = p_{\alpha i}\, q_{\alpha, i} - H(p, q, \boldsymbol{x}). \tag{3.62}$$

Let

$$\vec{A} = \left(\begin{array}{c|c} \vec{O}_1 & \vec{O}_1 \ldots \vec{O}_1 \\ \hline \vec{e}_1 \boldsymbol{E} & \\ \vdots & \vec{O}_2 \\ \vec{e}_n \boldsymbol{E} & \end{array} \right); \tag{3.63}$$

then

$$p_{\alpha i}\, q_{\alpha, i} = \{R\}' \vec{A} \bullet \vec{\nabla}\{R\}$$

where $\{R\}$ is the column matrix introduced in Section II. Thus,

$$\mathscr{L} = \tfrac{1}{2}\{R\}' \vec{\alpha} \bullet \vec{\nabla}\{\Psi\} - H(R, \boldsymbol{x}), \tag{3.64}$$

and equation (3.61) becomes

$$\vec{\nabla}\{R\}' \bullet \vec{\alpha} + H_{,\{R\}} = [0]$$

where

$$\vec{\alpha} = \vec{A} - \vec{A}'.$$

From the form of $\vec{A}$ given above, by comparison with equation (2.3), we conclude that $\vec{\alpha} = -\vec{I}$, and hence $\det|\vec{\alpha} \bullet \vec{\alpha}| \neq 0$, so that

$$-\vec{\nabla}\{R\}' \bullet \vec{I} + H_{,\{R\}} = [0],$$

which upon taking the transpose results in

$$\vec{I} \bullet \vec{\nabla}\{R\} + \{H_{,R}\} = \{0\}.$$

By Theorem 3.7, we may replace $\mathscr{L} = \{R\}' \vec{A} \bullet \vec{\nabla}\{R\} - H(R, \boldsymbol{x})$ by

$${}'\mathscr{L} = \tfrac{1}{2}\{R\}(\vec{A} - \vec{A}') \bullet \vec{\nabla}\{R\} - H(R, \boldsymbol{x}).$$

For this latter form of ${}'\mathscr{L}$ we see, upon using equation (3.63), that $p_{\alpha i}\, q_{\alpha, i}$ is variationally equivalent to

$$\tfrac{1}{2} p_{\alpha i}\, q_{\alpha, i} - \tfrac{1}{2} p_{\alpha i, i}\, q_{\alpha}$$

which implies that $p_{\alpha i}\, q_{\alpha, i}$ is variationally equivalent to $-p_{\alpha i, i}\, q_{\alpha}$.

The result is evident by Theorem 3.2 since

$$p_{\alpha i}\, q_{\alpha, i} = (p_{\alpha i}\, q_{\alpha})_{, i} - p_{\alpha i, i}\, q_{\alpha},$$

but is interesting in that it is arrived at solely from structural considerations without any consideration similar to Theorem 3.2.

As a further illustration, we apply the general method of alpha systems to the problem of Hamiltonian systems with constraints of the form

$$F_{\lambda \xi i} R_{\xi, i} - f_{\lambda}(R, \boldsymbol{x}) = 0, \qquad \lambda = 1, \ldots, k,$$

where $F_{\xi \lambda i}$ are constants. Rewriting the constraint equations, we have

$$\vec{F}_{\lambda} \bullet \vec{\nabla}\{R\} - f_{\lambda}(R, \boldsymbol{x}) = 0$$

where $\vec{F}_{\lambda}$ is a row vectrix defined by

$$\vec{F}_{\lambda} = [F_{\lambda 1 i} \vec{e}_i \;\; F_{\lambda 2 i} \vec{e}_i \; \ldots].$$

Multiplying the equations of constraint by undetermined multipliers η_λ and adding to the '$\mathscr{L}$ for the Hamiltonian system gives

$$\mathscr{L} = \tfrac{1}{2}\{\Psi\}'\vec{\alpha} \bullet \vec{\nabla}\{\Psi\} - H,$$

$$H = H(R, \boldsymbol{x}) + \eta_\lambda f_\lambda(R, \boldsymbol{x}),$$

$$\{\Psi\} = \begin{Bmatrix} \{R\} \\ \{\eta\} \end{Bmatrix},$$

$$\vec{\boldsymbol{J}} = \begin{Bmatrix} \vec{\boldsymbol{F}}_1 \\ \vdots \\ \vec{\boldsymbol{F}}_k \end{Bmatrix},$$

$$\vec{\alpha} = \left(\begin{array}{c|c} -\vec{\boldsymbol{I}} & -\vec{\boldsymbol{J}}' \\ \hline \vec{\boldsymbol{J}}' & \vec{\boldsymbol{O}} \end{array} \right),$$

from which we may proceed by well established methods to obtain the system equations in the presence of the given constraints.

Section IV. Properties of Canonical Maps

Group Properties

The first properties we shall establish are the group properties of the various types of canonical maps considered in Section II. These group properties are completely delineated by the following theorem and its corollaries:

***Theorem 4.1.** *The collection of all canonical maps form a group. In particular, if T_1 and T_2 are any two canonical maps with associated Jacobian matrices $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$, multipliers μ_1 and μ_2, and remainder functions U_1 and U_2 respectively, and $\boldsymbol{M}, \mu, U$ are the associated Jacobian matrix, multiplier and remainder function respectively of the canonical map obtained by applying T_2 after T_1, then the combinatorial laws for $\boldsymbol{M}, \mu, U$ are*

$$\boldsymbol{M} = \boldsymbol{M}_2 \boldsymbol{M}_1, \tag{4.1}$$

$$\mu = \mu_1 \mu_2, \tag{4.2}$$

$$U = U_2 + \mu_2 U_1. \tag{4.3}$$

Proof. By Theorem 2.2, the collection of all canonical maps are those of the collection $\mathscr{T}$ for which there exist nonzero constants μ such that the associated Jacobian matrices, $\boldsymbol{M}$, of the transformations satisfy the vectrix equations

$$\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} = \frac{1}{\mu}\vec{\boldsymbol{I}}.$$

For clarity, we note

$$\{R\} \underset{\mu_1 U_1}{\overset{T_1:\boldsymbol{M}}{\longleftrightarrow}} \{R_1\} \underset{\mu_2 U_2}{\overset{T_2:\boldsymbol{M}_2}{\longleftrightarrow}} \{R_2\},$$

$$\{R\} \underset{\mu U}{\overset{T:\boldsymbol{M}}{\longleftrightarrow}} \{\widetilde{R}_2\},$$

$$\{R_1\} = \{R_1(R, \boldsymbol{x})\}; \quad \{R_2\} = \{R_2(R_1, \boldsymbol{x})\},$$

$$\{\widetilde{R}_2\} = \{\widetilde{R}_2(R, \boldsymbol{x})\} = \{R_2(R_1(R, \boldsymbol{x}), \boldsymbol{x})\}$$

where the $\{R\}$'s are the $N(n+1)$-element column matrices introduced in Section II. It follows immediately, from the properties of matrices, that $\boldsymbol{M}=\boldsymbol{M}_2\boldsymbol{M}_1$. Thus $\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M}=\boldsymbol{M}_1'\boldsymbol{M}_2'\vec{\boldsymbol{I}}\boldsymbol{M}_2\boldsymbol{M}_1$. But $\boldsymbol{M}_1'\vec{\boldsymbol{I}}\boldsymbol{M}_1=\frac{1}{\mu_1}\vec{\boldsymbol{I}}$ and $\boldsymbol{M}_2'\vec{\boldsymbol{I}}\boldsymbol{M}_2=\frac{1}{\mu_2}\vec{\boldsymbol{I}}$, since both T_1 and T_2 are assumed to be canonical; hence $\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M}=\frac{1}{\mu_1\mu_2}\vec{\boldsymbol{I}}$. We may thus conclude that there exists a multiplier $\mu=\mu_1\mu_2$ such that $\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M}=\frac{1}{\mu}\vec{\boldsymbol{I}}$, and hence $\boldsymbol{M}$ is a generalized symplectic matrix of multiplier μ. Thus, the product of any two canonical maps is a canonical map (product interpreted in terms of the successive application of canonical maps). From equation (2.50), the remainder function is defined by $\{U_{,\tilde{R}_2}\}=\vec{\boldsymbol{I}}\bullet\underset{0}{\vec{\nabla}}\{R_2\}$. Now $\vec{\boldsymbol{I}}\bullet\underset{0}{\vec{\nabla}}\{\tilde{R}_2\}=\vec{\boldsymbol{I}}\bullet\underset{0}{\vec{\nabla}}\{R_2\}+\vec{\boldsymbol{I}}\{R_2\}_{,\{R_1\}}\bullet\underset{0}{\vec{\nabla}}\{R_1\}$ which, since $\{R_2\}_{,\{R_1\}}=\boldsymbol{M}_2^{-1}$, gives $\vec{\boldsymbol{I}}\bullet\underset{0}{\vec{\nabla}}\{\tilde{R}_2\}=\vec{\boldsymbol{I}}\bullet\underset{0}{\vec{\nabla}}\{R_2\}+\vec{\boldsymbol{I}}\boldsymbol{M}_2^{-1}\bullet\underset{0}{\vec{\nabla}}\{R_1\}$ $=\{U_{2,R_2}\}+\vec{\boldsymbol{I}}\boldsymbol{M}_2^{-1}\bullet\underset{0}{\vec{\nabla}}\{R_1\}$ by definition of U_2. However, $\boldsymbol{M}_2^{-1\prime}\vec{\boldsymbol{I}}\boldsymbol{M}_2^{-1}=\mu_2\vec{\boldsymbol{I}}$, so that $\vec{\boldsymbol{I}}\boldsymbol{M}_2^{-1}=\mu_2\boldsymbol{M}_2'\vec{\boldsymbol{I}}$, and hence

$$\begin{aligned}\vec{\boldsymbol{I}}\bullet\underset{0}{\vec{\nabla}}\{R_2\}&=\{U_{2,R_2}\}+\mu_2\boldsymbol{M}_2'\vec{\boldsymbol{I}}\bullet\underset{0}{\vec{\nabla}}\{R_1\},\\&=\{U_{2,R_2}\}+\mu_2\boldsymbol{M}_2'\{U_{1,R_1}\},\\&=\{U_{2,R_2}\}+\mu_2\{U_{1,R_2}\}.\end{aligned}$$

Thus, the composite remainder function is given by $U=U_2+\mu_2U_1$ to within an additive function of $\boldsymbol{x}$ which may be deleted with no loss of generality. Equations (4.1) through (4.3) are thus established. It now remains to prove that the collection of all canonical maps contains the inverse of every element and the identity element. Taking $\boldsymbol{M}$ to be the identity matrix, $U=0$ and $\mu=1$ result in $\{R\}=\{\tilde{R}\}$ and $K=H$ by equation (2.49) and hence represents the identity map. From equations (4.1) through (4.3) it is evident that if $\boldsymbol{M}_1$ is the inverse of $\boldsymbol{M}_2$, then $\boldsymbol{M}_2\boldsymbol{M}_1=\boldsymbol{E}^*$ whence $\mu_2\mu_1=1$ and $0\equiv U_2+\mu_2U_1$, so that $\boldsymbol{M}_1=\boldsymbol{M}_2^{-1}$, $\mu_2=\frac{1}{\mu_1}$, $U_2=-\mu_2U_1$, and hence the inverse of every canonical map is a canonical map. Q.E.D.

The group of all canonical maps, established by the above theorem, is a generalization of what is called the symplectic group[13] for the case $n=1$, and will be referred to as the **generalized symplectic group**. It is the generalized symplectic group which is the **invariance group for systems of generalized Hamiltonian equations**. The system of equations (2.1) or their vectrix equivalents, equations (2.12), are invariant in form under this group for all bases functions.

The following corollaries to Theorem 4.1 are immediate upon noting that all the subcollections of canonical maps considered contain the identity element.

***Corollary.** *The collection of all uniform canonical maps forms a group, the* **uniform symplectic group.**

***Corollary.** *The collection of all basic canonical maps forms a group, the* **basic symplectic group.**

---

[13] Siegel, C. L.: Vorlesungen über Himmelsmechanik, p. 9. Berlin-Göttingen-Heidelberg: Springer 1956.

*Corollary. *The collection of all absolutely uniform canonical maps forms a group, the* **absolutely uniform symplectic group.**

Consider the subgroup of the generalized symplectic group which are linear and homogeneous in (p, q) (that is, for which $\{\widetilde{R}\} = \boldsymbol{M}^{-1}\{R\}$ where $\boldsymbol{M} = \boldsymbol{M}(\boldsymbol{x})$ is a nonsingular matrix for all $\boldsymbol{x}$ in $\mathscr{D}_n^*$). For this subgroup we have $\boldsymbol{M}^{-1\prime}\vec{\boldsymbol{I}}\boldsymbol{M}^{-1} = \mu\vec{\boldsymbol{I}}$ and $\underset{0}{\vec{\nabla}}\{\widetilde{R}\} = (\underset{0}{\vec{\nabla}}\boldsymbol{M})\{R\}$. Now, by equations (2.48)

$$\{U_{,\widetilde{R}}\} = \vec{\boldsymbol{I}} \bullet \underset{0}{\vec{\nabla}}\{\widetilde{R}\} = \vec{\boldsymbol{I}} \bullet (\underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1})\{R\}$$
$$= \vec{\boldsymbol{I}} \bullet (\underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1})\boldsymbol{M}\{\widetilde{R}\},$$

so that the remainder function is given by the quadratic form

$$U = \tfrac{1}{2}\{\widetilde{R}\}\vec{\boldsymbol{I}} \bullet (\underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1})\boldsymbol{M}\{\widetilde{R}\},$$

and hence, by equation (2.49),

$$\mathrm{K}(\widetilde{R}, \boldsymbol{x}) = \mu H(\boldsymbol{M}\{\widetilde{R}\}, \boldsymbol{x}) - \tfrac{1}{2}\{\widetilde{R}\}\vec{\boldsymbol{I}} \bullet (\underset{0}{\vec{\nabla}}\boldsymbol{M}^{-1})\boldsymbol{M}\{\widetilde{R}\}.$$

As an example, consider the matrix $\boldsymbol{M}^{-1}$ given by repeating along the principal diagonal an orthogonal $N \times N$ matrix $\boldsymbol{v}(\boldsymbol{x})$, that is,

$$\boldsymbol{M}^{-1} = \begin{pmatrix} \boldsymbol{v} & 0 & \dots & \\ 0 & \boldsymbol{v} & & \\ \vdots & & \ddots & \vdots \\ & \dots & & \boldsymbol{v} \end{pmatrix} \qquad \begin{aligned} \boldsymbol{v}' &= \boldsymbol{v}^{-1} \\ \boldsymbol{M}' &= \boldsymbol{M}^{-1}. \end{aligned}$$

It is easily verified that $\boldsymbol{M}^{-1\prime}\vec{\boldsymbol{I}}\boldsymbol{M}^{-1} = \vec{\boldsymbol{I}}$, so that such a map is an element of the basic symplectic group and hence will generate a canonical map. For this case, the remainder function is given by

$$U = \tfrac{1}{2}\left([Q_\alpha](\partial_i \boldsymbol{v})'\boldsymbol{v}\{P_{\alpha i}\} - [P_{\alpha i}](\partial_i \boldsymbol{v})'\boldsymbol{v}\{Q_\alpha\}\right).$$

There is a "factorization" of the generalized symplectic group by which we may represent any canonical map as the product of a homogeneous linear dilatation and a basic map.

*Theorem 4.2. *The generalized symplectic groups admit a factorization into the basic group and the group of all homogeneous linear dilatations (all canonical maps whose associated Jacobian matrix is of the form $\frac{1}{\sqrt{a}}\boldsymbol{E}^*$ for any a bounded away from zero), in the sense that if $\boldsymbol{M}$ is any symplectic matrix of multiplier μ then $\boldsymbol{M} = \boldsymbol{M}_1\boldsymbol{M}_2$ where $\boldsymbol{M}_1 = \frac{1}{\sqrt{\mu}}\boldsymbol{E}^*$ and $\boldsymbol{M}_2 = \boldsymbol{M}_1^{-1}\boldsymbol{M}$ is basic.*

Proof. First we note that $\boldsymbol{M}_1 = \frac{1}{\sqrt{\mu}}\boldsymbol{E}^*$ is an element of the generalized symplectic group with multiplier μ for $\boldsymbol{M}_1'\vec{\boldsymbol{I}}\boldsymbol{M}_1 = \frac{1}{\mu}\vec{\boldsymbol{I}}$ and that the collection of all such homogeneous linear dilatations with associated Jacobian matrix $\frac{1}{\sqrt{\alpha}}\boldsymbol{E}^*$ form a group. By Theorem 4.1, $\boldsymbol{M}_2$ is an element of the generalized symplectic group

since $\boldsymbol{M}$ and $\boldsymbol{M}_1$ are elements, and $\boldsymbol{M}_2$ has multiplier $\mu_2=\mu/\mu_1=\mu/\mu=1$ by equation (4.2) since $\boldsymbol{M}_1^{-1}$ has multiplier $1/\mu$. Thus, $\boldsymbol{M}_2$ is an element of the basic symplectic group. Q.E.D.

If we consider the $N(n+1)$-element column matrix $\{R\}$ as an ordered collection of $N(n+1)$ numbers, then we may represent $\{R\}$ as a point in a Euclidean space I of $N(n+1)$ dimensions by making a one-to-one correspondence between the values of the elements of $\{R\}$ and the coordinates in I. Such a space I will be refered to as a **generalized phase space.** From the definition of the uniform symplectic group, we see that it is a group of isomorphism of I. If we fix the value of $\boldsymbol{x}$ in $\mathscr{D}_n^*$, then the generalized symplectic group is also a group of isomorphisms of I. Thus, the generalized symplectic group represents, in general, an n-parameters collection of isomorphisms of I where the n parameters are the $\boldsymbol{x}$ whose domain is $\mathscr{D}_n^*$.

Let H be an element of $\mathfrak{H}\ominus\mathfrak{H}_0$, so that there exists at least one solution to equations (2.12), $\{R\}=\{R(\boldsymbol{x})\}$. We may interpret $\{R(\boldsymbol{x})\}$ as an n-dimensional hypersurface in I. Thus, the Hamiltonian equations may be considered as generating n-dimensional hypersurfaces in I. When $\{R\}$ is considered as a function $\boldsymbol{x}$, then I will be refered to as the **solution space** I_s (that is, in the solution space the $\{R\}$ are assumed to be functions of $\boldsymbol{x}$ and hence are not independent).

Lagrange Bracket Vectrices

Many of the results of the classic theory stem from considerations of certain bracket expressions refered to as Lagrange and Poisson brackets. The question naturally arises as to the form and properties of the generalizations of the classic bracket expressions as determined by the properties of the generalized symplectic group. The forms we use for definition are similar to those used by WINTNER[14] although, as might be expected, we must use the structure vectrix of the system involved.

Definition.* By the **Lagrange bracket n-vectrix, $\vec{L}(u, v; R)$, is meant the vectrix collection of bilinear polar forms defined over the generalized phase space I by

$$\vec{L}(u, v; R) \stackrel{\text{def}}{=} \{R\}'_{,u}\,\vec{\boldsymbol{I}}\{R\}_{,v} \tag{4.4}$$

where $\{R\}$ is the $N(n+1)$-element column matrix defined in Section II and (u, v) are any two of a collection of parameters which define a hypersurface in I.

The Lagrange bracket vectrix so defined is the generalization, for the case $n>1$, of the bracket expressions introduced by LAGRANGE[15] in 1808. We have already encountered the components of the Lagrange bracket vectrix in Section III, where they were used to facilitate the alternate development of canonical maps from the standpoint of the calculus of variations.

Theorem 4.3. *The Lagrange bracket vectrix exhibits (for any (u,v) satisfying the conditions of the above definition) the following properties:*

(a) $$\vec{L}(u, v; R) = \{i\,|\,v, u; p\,q\}\,\vec{e}_i \tag{4.5}$$

[14] WINTNER, A.: The Analytical Foundations of Celestial Mechanics.
[15] LAGRANGE, J. L.: Mém. de l'Inst. de France 1808.

where $\{i|\, u, v; p\, \mathrm{q}\}$ was defined in Theorem 3.1 to be $p_{\alpha i,u}\mathrm{q}_{\alpha,v} - p_{\alpha i,v}\mathrm{q}_{\alpha,u}$,

$$\text{*(b)} \qquad \vec{L}(u, v; R) = -\vec{L}(v, u; R), \tag{4.6}$$

$$\text{(c)} \qquad \vec{L}(\mathrm{q}_\alpha, \mathrm{q}_\beta; R) = 0, \tag{4.7}$$

$$\text{(d)} \qquad \vec{L}(p_{\alpha i}, p_{\beta j}; R) = 0, \tag{4.8}$$

$$\text{(e)} \qquad \vec{L}(\mathrm{q}_\beta, p_{\alpha i}; R) = \delta_{\alpha\beta}\,\vec{e}_i, \tag{4.9}$$

$$\text{(f)} \qquad \vec{L}(\mathrm{q}_\beta, p_{\gamma j}; R) \bullet \vec{L}(\mathrm{q}_\psi, p_{\lambda k}; R) = \delta_{\beta\gamma}\,\delta_{\psi\lambda}\,\delta_{jk}, \tag{4.10}$$

$$\text{(g)} \qquad \begin{aligned} \vec{L}(u, v; R) = [&\vec{L}(\mathrm{q}_\beta, u; R) \bullet \vec{L}(p_{\beta j}, v; R) - \\ &- \vec{L}(\mathrm{q}_\beta, v; R) \bullet \vec{L}(p_{\beta j}, u; R)]\,\vec{e}_i. \end{aligned} \tag{4.11}$$

*(h) *Let T be any element of the generalized symplectic group of multiplier μ, $T\colon \{R\} \to \{\tilde{R}\}$; then*

$$\vec{L}(u, v; R) = \frac{1}{\mu}\,\vec{L}(u, v; \tilde{R}), \tag{4.12}$$

and conversely if (4.12) follows for any (u, v).

Proof. Part (a) results directly from an expansion of equation (4.4). Part (b) results from noting that $\vec{\boldsymbol{I}}$ is antisymmetric (that is,

$$\vec{L}(v, u) = \{R\}'_{,v}\,\vec{\boldsymbol{I}}\{R\}_{,u} = \{R\}'_{,u}\,\vec{\boldsymbol{I}}\{R\}_{,v} = -\vec{L}(u, v)$$

since $\vec{\boldsymbol{I}} = -\vec{\boldsymbol{I}}'$). From (a) we have

$$\vec{L}(\mathrm{q}_\gamma, \mathrm{q}_\beta; R) = \{i|\, \mathrm{q}_\beta, \mathrm{q}_\gamma; p\,\mathrm{q}\}\,\vec{e}_i = (p_{\alpha i,\mathrm{q}_\beta}\mathrm{q}_{\alpha,\mathrm{q}_\lambda} - p_{\alpha i,\mathrm{q}_\gamma}\mathrm{q}_{\alpha,\mathrm{q}_\beta})\,\vec{e}_i = 0$$

since p and q are independent in I. Thus (c) is proved. Proceeding in a similar manner, (d) may be established immediately. Using (a), we have

$$\begin{aligned} \vec{L}(\mathrm{q}_\beta, p_{\gamma j}; R) &= \{i|\, p_{\gamma j}, \mathrm{q}_\beta; p\,\mathrm{q}\}\,\vec{e}_i, \\ &= (p_{\alpha i,p_{\gamma j}}\mathrm{q}_{\alpha,\mathrm{q}_\beta} - p_{\alpha i,\mathrm{q}_\beta}\mathrm{q}_{\alpha,p_{\gamma j}})\,\vec{e}_i, \\ &= (p_{\alpha i,p_{\gamma j}}\mathrm{q}_{\alpha,\mathrm{q}_\beta})\,\vec{e}_i = \delta_{\beta\gamma}\,\vec{e}_j, \end{aligned}$$

since p and q are independent in I, and hence $\vec{L}(\mathrm{q}_\beta, p_{\gamma j}; R) = \delta_{\beta\gamma}\vec{e}_i$. Thus (e) is established. (f) follows immediately from (e). Noting that p and q are independent in I, from (a) we have

$$\begin{aligned} \vec{L}(\mathrm{q}_\beta, v; R) &= p_{\beta i,v}\,\vec{e}_i, \\ \vec{L}(p_{\beta j}, v; R) &= -\,\mathrm{q}_{\beta,v}\,\vec{e}_j. \end{aligned}$$

Hence

$$\begin{aligned} &\vec{L}(\mathrm{q}_\beta, u; R) \bullet \vec{L}(p_{\beta j}, v; R) - \vec{L}(\mathrm{q}_\beta, v; R) \bullet \vec{L}(p_{\beta j}, u; R) \\ &\quad = -\,p_{\beta i,u}\mathrm{q}_{\beta,v}\,\vec{e}_i \bullet \vec{e}_j + p_{\beta i,v}\mathrm{q}_{\beta,u}\,\vec{e}_i \bullet \vec{e}_j, \\ &\quad = \mathrm{q}_{\beta,u}p_{\beta j,v} - \mathrm{q}_{\beta,v}p_{\beta j,u}, \end{aligned}$$

which by (a) is just the j^{th} component of the Lagrange bracket vectrix $\vec{L}(u, v; R)$. Thus (g) is established. By definition $\vec{L}(u, v; R) = \{R\}'_{,u}\vec{\boldsymbol{I}}\{R\}_{,v}$. Now $\{R\}_{,u} = \boldsymbol{M}\{\tilde{R}\}_{,u}$, where $\boldsymbol{M}$ is the associated Jacobian matrix of the canonical map considered, so that $\vec{L}(u, v; R) = \{R\}'_{,u}\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M}\{R\}_{,v} = \frac{1}{\mu}\{R\}'_{,u}\boldsymbol{I}\{R\}_{,v}$, by Theorem 2.1. Since the last expression is equal to $\frac{1}{\mu}\vec{L}(u, v; \tilde{R})$ by definition, the result follows. On the other hand, if $\vec{L}(u, v; \tilde{R}) = \mu\vec{L}(u, v; R)$, then $\{\tilde{R}\}'_{,u}\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M}\{R\}_{,v} = \frac{1}{\mu}\{\tilde{R}\}'_{,u}\vec{\boldsymbol{I}}\{\tilde{R}\}_{,v}$ so that $\{\tilde{R}\}'_{,u}\left[\boldsymbol{M}'\boldsymbol{I}\boldsymbol{M} - \frac{1}{\mu}\vec{\boldsymbol{I}}\right]\{\tilde{R}\}_{,v} = 0$. If this is to be true for all u and v, then certainly it must hold for any (u, v) elements of $\{\tilde{R}\}$ and hence for $\{\tilde{R}\}$ so that

$$(\{\tilde{R}\}_{,\{\tilde{R}\}})'\left[\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} - \frac{1}{\mu}\vec{\boldsymbol{I}}\right](\{\vec{R}\}_{,\{\tilde{R}\}}) = \boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} - \frac{1}{\mu}\vec{\boldsymbol{I}} = \vec{\boldsymbol{O}}. \quad \text{Q.E.D.}$$

Part (h) of the last theorem allows us to calculate the Lagrange bracket vectrix for any $\{\tilde{R}\}$ obtained from a known $\{R\}$ by a canonical map simply through multiplication by the reciprocal of the multiplier of the map considered. We may thus drop the superfluous R in writing $\vec{L}(u, v; R)$ $\left(\text{that is, } \vec{L}(u, v) \overset{\text{def}}{=} \vec{L}(u, v; R)\right)$.

The Lagrange bracket vectrix possesses an additional property under the generalized symplectic group which is similar to that of the integral invariants of Poincaré[16].

Definition.* A transformation is said to admit an **absolute integral, $\mathscr{I}$, **of multiplier** μ **and order** W if and only if the integral of $\mathscr{I}$ over every W-dimensional compact point set in the space of functions over which the transformation operates changes only by multiplication by $1/\mu$ when the points of the space undergo the transformation considered. If $\mu = 1$, $\mathscr{I}$ is said to be an **absolute integral invariant of order** $\boldsymbol{W}$.

***Theorem 4.4.** *Every element of the generalized symplectic group of multiplier μ admits the absolute integrals*

$$\vec{L}(u, v)$$

of multiplier μ and order 2. In addition, the basic symplectic group admits the absolute integral invariants

$$\vec{L}(u, v)$$

of order 2.

Proof. From Theorem 4.3 (h) we have, for any element of the generalized symplectic group of multiplier μ, $\vec{L}(u, v; R) = \frac{1}{\mu}\vec{L}(u, v; \tilde{R})$ for any (u, v). If S is any compact two-dimensional point set in the generalized phase space I, then it may be characterized by a two-parameter system (u, v). Hence

$$\int_S \vec{L}(u, v; R)\, du\, dv = \int_S \frac{1}{\mu}\vec{L}(u, v; \tilde{R})\, du\, dv$$

for all S in I. The second statement follows immediately upon noting that the basic symplectic group has only unity for the multiplier of its elements. Q.E.D.

[16] Poincaré, H.: Acta Math. **13** (1890).

Lagrange-S Bracket Vectrices

There is another vectrix bracket whose form is similar to the Lagrange bracket vectrix, but which arises only in the case of more than one independent variable.

***Definition.* By the **Lagrange-S bracket vectrix** $\overrightarrow{LS}(x^i, x^j; R)$ is meant the vectrix collection of bilinear polar forms defined over the solution space I_s by

$$\overrightarrow{LS}(x^i, x^j; R) \stackrel{\text{def}}{=} \{R\}'_{,i}\vec{\boldsymbol{I}}\{R\}_{,j}. \tag{4.13}$$

From this definition, we have

***Theorem 4.5.** *Let T be an element of the generalized symplectic group with associated Jacobian matrix M and multiplier μ; then*

$$\begin{aligned}\overrightarrow{LS}(x^i, x^j; R) = \frac{1}{\mu}\,[\overrightarrow{LS}(x^i, x^j; \tilde{R}) + (\partial_i\{\tilde{R}\})'\vec{\boldsymbol{I}}\partial_j\{\tilde{R}\} -\\ -\{\tilde{R}\}'_{,i}\vec{\boldsymbol{I}}\partial_j\{\tilde{R}\} - (\partial_i\{\tilde{R}\})'\vec{\boldsymbol{I}}\{\tilde{R}\}_{,j}].\end{aligned} \tag{4.14}$$

Proof. We have seen in Section II that $\{R\}_{,i} = \boldsymbol{M}\{\tilde{R}\}_{,i} - \boldsymbol{M}\partial_i\{\tilde{R}\}$ under T. Thus

$$\begin{aligned}\overrightarrow{LS}(x^i, x^j; R) &= \{R\}'_{,i}\vec{\boldsymbol{I}}\{R\}_{,j},\\ &= (\{\tilde{R}\}'_{,i}\boldsymbol{M}' - \partial_i\{\tilde{R}\}'\boldsymbol{M}')\,\vec{\boldsymbol{I}}(\boldsymbol{M}\{\tilde{R}\}_{,j} - \boldsymbol{M}\partial_j\{\tilde{R}\}),\\ &= (\{\tilde{R}\}'_{,i} - \partial_i\{\tilde{R}\}')\,\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M}(\{\tilde{R}\}_{,j} - \partial_j\{\tilde{R}\}),\\ &= \frac{1}{\mu}\,(\{\tilde{R}\}'_{,i} - \partial_i\{\tilde{R}\}')\,\vec{\boldsymbol{I}}(\{\tilde{R}\}_{,j} - \partial_j\{\tilde{R}\}),\\ &= \frac{1}{\mu}\,[\{\tilde{R}\}'_{,i}\vec{\boldsymbol{I}}\{\tilde{R}\}_{,j} - (\partial_i\{\tilde{R}\})'\vec{\boldsymbol{I}}\{\tilde{R}\}_{,j} -\\ &\quad - \{\tilde{R}\}'_{,i}\vec{\boldsymbol{I}}\partial_j\{\tilde{R}\} + (\partial_i\{\tilde{R}\})'\vec{\boldsymbol{I}}\partial_j\{\tilde{R}\}].\end{aligned}$$

But $\overrightarrow{LS}(x^i, x^j; \tilde{R}) = \{\tilde{R}\}'_{,i}\,\vec{\boldsymbol{I}}\{\tilde{R}\}_{,j}$ by equation (4.13), whence the result follows. Q.E.D.

***Corollary.** *Under a uniform canonical map, the Lagrange-S bracket vectrix changes only by multiplication by the reciprocal of the multiplier of the map considered.*

Proof. The result follows immediately from Theorem 4.5 upon noting that $\partial_i\{\tilde{R}\} \equiv 0$ for all elements of the uniform symplectic group, by definition. Q.E.D.

***Theorem 4.6.** *Every element of the uniform symplectic group of multiplier μ admits the absolute integrals $\overrightarrow{LS}(x^i, x^j; R)$ of multiplier μ and weight 2. In addition, the absolutely uniform symplectic group admits the absolute integral invariants $\overrightarrow{LS}(x^i, x^j; R)$ of weight 2.*

Proof. The proof of this theorem follows from the above corollary by the same reasoning as in the proof of Theorem 4.4. Q.E.D.

Let

$$\overrightarrow{LS}(\otimes\boldsymbol{x}, x^i; R) \stackrel{\text{def}}{=} \vec{V}\{R\}'\bullet\vec{\boldsymbol{I}}\{R\}_{,i} \tag{4.15a}$$

and

$$\overrightarrow{LS}(x^i, \otimes\boldsymbol{x}; R) \stackrel{\text{def}}{=} \{R\}'_{,i}\vec{\boldsymbol{I}}\bullet\vec{V}\{R\}. \tag{4.15b}$$

From equation (2.12) we have $\vec{\boldsymbol{I}}\bullet\vec{V}\{R\} = -\{H_{,R}\}$, so that $\vec{V}\{R\}'\bullet\vec{\boldsymbol{I}} = \{H_{,R}\}'$. Thus equation (4.15a) gives $\overrightarrow{LS}(\otimes\boldsymbol{x}, x^i; R) = \{H_{,R}\}'\,\{R\}_{,i}$ if $\{R\}$ is an n-dimen-

sional solution manifold in I_s which constitutes a solution to equation (2.12). Now

$$H_{,i} = \{H_{,R}\}' \{R\}_{,i} + \partial_i H = \overrightarrow{LS}(\otimes \boldsymbol{x}, x^i; R) + \partial_i H$$

or

$$\vec{V} H = \overrightarrow{LS}(\otimes \boldsymbol{x}, \boldsymbol{x}; R) + \underset{0}{\vec{V}} H.$$

We have thus proved the following

***Theorem 4.7.** *Let $\mathfrak{M}$ be any n-dimensional hypersurface in I_s which constitutes a solution manifold of the generalized Hamiltonian equations* (2.12); *then*

$$\vec{V} H = \overrightarrow{LS}(\otimes \boldsymbol{x}, \boldsymbol{x}; R) + \underset{0}{\vec{V}} H \tag{4.16}$$

is an identity over $\mathfrak{M}$, where

$$\overrightarrow{LS}(\otimes \boldsymbol{x}, \boldsymbol{x}; R) \overset{\text{def}}{=} (\vec{V}\{R\}' \bullet \vec{\boldsymbol{I}})\, \vec{V}\{R\}.$$

***Theorem 4.8.** *Let $\mathfrak{M}$ be any n-dimensional hypersurface of I_s which constitutes a solution manifold to equations* (2.12) *for $\boldsymbol{x}$ in a given $\mathscr{D}_n^*$, where both H and the solution $\{R\}$ are of class C^2; then*

$$\begin{aligned} \vec{V} \bullet \overrightarrow{LS}(x^i, x^j; R) &= \partial_i H_{,j} - \partial_j H_{,i} \\ &= \partial_i \overrightarrow{LS}(\otimes \boldsymbol{x}, x^j; R) - \partial_j \overrightarrow{LS}(\otimes \boldsymbol{x}, x^i; R) \end{aligned} \tag{4.17}$$

is an identity in $\mathfrak{M}$ for all $\mathfrak{M}$.

Proof. From the definition of $\overrightarrow{LS}(x^i, x^j; R)$, we have

$$\begin{aligned} \vec{V} \bullet \overrightarrow{LS}(x^i, x^j; R) &= \vec{V} \bullet (\{R\}'_{,i} \vec{\boldsymbol{I}} \{R\}_{,j}) \\ &= [\vec{V}(\{R\}'_{,i}) \bullet \vec{\boldsymbol{I}}]\, \{R\}_{,j} + \{R\}'_{,i}\, [\vec{\boldsymbol{I}} \bullet \vec{V}(\{R\}_{,j})] \end{aligned}$$

since $\vec{\boldsymbol{I}}$ is a constant vectrix. In $\mathfrak{M}$, $\vec{\boldsymbol{I}} \bullet \vec{V}\{R\} + \{H_{,R}\} = \{0\}$ so that $\vec{\boldsymbol{I}} \bullet [\vec{V}\{R\}]_{,i} = -\{H_{,R}\}_{,i} = \vec{\boldsymbol{I}} \bullet \vec{V}(\{R\}_{,i})$ since $\{R\}$ was assumed to be of class C^2. Thus

$$\vec{V} \bullet \overrightarrow{LS}(x^i, x^j; R) = \{H_{,R}\}'_{,i} \{R\}_{,j} - \{R\}'_{,i} \{H_{,R}\}_{,j}.$$

However

$$(H_{,i})_{,j} = \{H_{,R}\}_{,i} \{R\}_{,j} + \partial_j H_{,i},$$

so that

$$\vec{V} \bullet \overrightarrow{LS}(x^i, x^j; R) = H_{,ij} - \partial_j H_{,i} - H_{,ji} + \partial_i H_{,j} = \partial_i H_{,j} - \partial_j H_{,i}$$

since both H and $\{R\}$ were assumed to be of class C^2. This result may also be written as

$$\vec{V} \bullet \overrightarrow{LS}(x^i, x^j; R) = \partial_i \overrightarrow{LS}(\otimes \boldsymbol{x}, x^j; R) - \partial_j \overrightarrow{LS}(\otimes \boldsymbol{x}, x^i; R)$$

since $H_{,j} = \overrightarrow{LS}(\otimes \boldsymbol{x}, x^j) + \partial_j H$, and thus

$$\partial_i H_{,j} - \partial_j H_{,i} = \partial_i \overrightarrow{LS}(\otimes \boldsymbol{x}, x^j; R) - \partial_j \overrightarrow{LS}(\otimes \boldsymbol{x}, x^i; R). \qquad \text{Q.E.D.}$$

Integrating equation (4.17) over $\mathscr{D}_n^*$ and using equation (1.16), we obtain

$$\int_{\mathscr{D}_n^* \ominus \mathscr{D}_n} \overrightarrow{LS}(x^i, x^j; R) \bullet \vec{N}\, dS = \int_{\mathscr{D}_n^*} \left(\partial_i \overrightarrow{LS}(\otimes \boldsymbol{x}, x^j; R) - \partial_j \overrightarrow{LS}(\otimes \boldsymbol{x}, x^i; R)\right) dv_n. \tag{4.18}$$

If $\partial_i H_{,j} - \partial_j H_{,i}$ is zero in $\mathfrak{M}$, then we have from equation (4.17)

$$0 = \int_{\mathcal{D}_n^*} \vec{V} \bullet \overrightarrow{LS}(x^i, x^j; R)\, dv_n,$$

so that $\overrightarrow{LS}(x^i, x^j; R)$ is conserved under equations (2.12) if $\partial_i H_{,j} - \partial_j H_{,i} = 0$ in $\mathfrak{M}$.

The converse of Theorem 4.8 is of interest in its own right.

*Theorem 4.9. *Let $\{R\}$ satisfy the equation $\vec{I} \bullet \vec{V}\{R\} + \{G(R, \boldsymbol{x})\} = \{0\}$ where $\{G(R, \boldsymbol{x})\}$ is an $N(n+1)$ element column matrix whose elements are class C^1, so that $\{R\}$ defines an n-dimensional hypersurface $\mathfrak{M}_G$ in I_s. If $\vec{V} \bullet \overrightarrow{LS}(x^i, x^j; R) = (\partial_i \{G\}')\{R\}_{,j} - \{R\}'_{,i}\, \partial_j \{G\}$ holds at all points in $\mathfrak{M}_G$, then there exists a function H such that $\{G\} = \{H_{,R}\}$ (that is, the system is Hamiltonian).*

Proof. $\vec{V} \bullet \overrightarrow{LS}(x^i, x^j; R) = (\vec{V}\{R\}'_{,i}) \bullet \vec{I}\{R\}_{,j} + \{R\}'_{,i}\vec{I} \bullet \vec{V}\{R\}_{,j}$. In $\mathfrak{M}$ we have $\vec{I} \bullet \vec{V}\{R\} = -\{G\}$ so that

$$\vec{V} \bullet \overrightarrow{LS}(x^i, x^j; R) = \{G\}'_{,i}\{R\}_{,j} - \{R\}'_{,i}\{G\}_{,j}.$$

But $\{G\}_{,i} = \{G\}_{,\{R\}}\{R\}_{,i} + \partial_i\{G\}$, and hence

$$\begin{aligned}\vec{V} \bullet \overrightarrow{LS}(x^i, x^j; R) &= [\{R\}'_{,i}(\{G\}_{,\{R\}})' + \partial_i\{G\}]\{R\}_{,j} - \{R\}'_{,i}[\{G\}_{,\{R\}}\{R\}_{,j} + \partial_j\{G\}] \\ &= \{R\}'_{,i}[(\{G\}_{,\{R\}})' - \{G\}_{,\{R\}}]\{R\}_{,j} + (\partial_i\{G\})'\{R\}_{,j} - \{R\}'_{,i}\,\partial_j\{G\}.\end{aligned}$$

Thus, the hypothesis can be satisfied only if

$$(\{G\}_{,\{R\}})' = \{G\}_{,\{R\}},$$

which is a necessary and sufficient condition, under the assumed continuity conditions, that there exist a function H such that $\{G\} = \{H_{,R}\}$. Q.E.D.

In Section VIII we show that any Hamiltonian system which is absolutely integrable is such that its solution may be represented locally by $\{R\} = \{R(\boldsymbol{x}, \{C\})\}$ where $\{C\}$ is an $N(n+1)$ element column matrix of constants. Using this result, we have the following theorem, which may be thought of as constituting a correspondence between the considerations of Lagrange bracket vectrices and Lagrange-S bracket vectrices.

*Theorem 4.10. *Let $\mathfrak{M}$ be any n-dimensional hypersurface of the generalized phase space I_s which constitutes a solution manifold of (2.12) for $\{R\}$ of class C^2; then $\vec{V} \bullet \vec{L}(\{C\}, \{C\}) = \boldsymbol{O}$ is an identity in $\mathfrak{M}$. Conversely, if $\{R\}$ is a solution to $\vec{I} \bullet \vec{V}\{R\} + \{G(R, \boldsymbol{x})\} = \{0\}$ where $\{G\}$ is of class C^1 and $\vec{V} \bullet \vec{L}(\{C\}, \{C\}) = \boldsymbol{O}$ at all points in $\mathfrak{M}$, then there exists a function H of class C^2 such that $\{G\} = \{H_{,R}\}$.*

Proof.

$$\begin{aligned}\vec{V} \bullet \vec{L}(\{C\}, \}C\}) &= \vec{V} \bullet [(\{R\}_{,\{C\}})'\vec{I}\{R\}_{,\{C\}}] \\ &= [\vec{V}(\{R\}_{,\{C\}})'] \bullet \vec{I}\{R\}_{,\{C\}} + (\{R\}_{,\{C\}})'\vec{I} \bullet \vec{V}(\{R\}_{,\{C\}}).\end{aligned}$$

Now $\vec{V}(\{R\}_{,\{C\}}) = (\vec{V}\{R\})_{,\{C\}}$, since $\{C\}$ and $\boldsymbol{x}$ are independent in $\mathfrak{M}$ so that $\vec{I} \bullet \vec{V}(\{R\}_{,\{C\}}) = \vec{I} \bullet (\vec{V}\{R\})_{,\{C\}} = -\{H_{,R}\}_{,\{C\}}$. From this point on, the proof proceeds in exactly the same way as in Theorems 4.8 and 4.9. Q.E.D.

Poisson Bracket Vectrices and their Lack of Invariance Under the Generalized Symplectic Group

Let u and v be two arbitrary functions of class C^1 defined over I.

*Definition**. By the **Poisson bracket vectrix** is meant the vectrix defined by the equation

$$\vec{P}(u, v; R) \stackrel{\text{def}}{=} u_{,\{R\}} \vec{\boldsymbol{I}}\{v_{,R}\}. \tag{4.19}$$

This vectrix is a direct generalization of the bracket expression first introduced by POISSON[17] for the case $n=1$. Although the form introduced by equation (4.19) will be shown to be non-invariant, its study is included for two reasons: first, the form of equation (4.19) admits direct commutation relations which do not depend on the dimension of the space $\mathscr{E}_n$ as seen from the antysymmetric character of the structure vectrix involved; and second, this form of bracket vectrix admits a direct extension of the Jacobi identity for the classic Poisson brackets.

The properties of the Poisson bracket vectrices are given by the following theorem:

Theorem 4.11. *The Poisson bracket vectors admit, for any* $(u, v, \ldots)$ *of class* C^1, *the following properties:*

(a) $$\vec{P}(u, v; R) = [i|\, v, u;\, R]\, \vec{e}_i$$

where $[i|\, u, v;\, R] = v_{,p_{\alpha i}} u_{,q_\alpha}, - v_{,q_\alpha} u_{,p_{\alpha i}},$

*(b) $$\vec{P}(u, v; R) = -\vec{P}(v, u; R),$$

(c) $$\vec{P}(p_{\beta j}, p_{\gamma k}; R) \equiv \vec{O},$$

(d) $$\vec{P}(\mathrm{q}_\beta, \mathrm{q}_\gamma; R) \equiv \vec{O},$$

(e) $$\vec{P}(\mathrm{q}_\beta, p_{\gamma j}; R) = \delta_{\beta\gamma} \vec{e}_j,$$

*(f) $$\vec{P}(u, u; R) = \vec{O},$$

*(g) $$\vec{P}(u+v, w; R) = \vec{P}(u, w; R) + \vec{P}(v, w; R),$$

(h) $$\vec{P}(u * v; R) = \vec{P}(u, w; R) * v + u * \vec{P}(v, w; R).$$

(i) *If* u, v *and* w *are any three functions of class* C^2 *defined over* I, *then*

$$\begin{aligned} &[i|\,[j|\, u, v; R], w; R] + [i|\,[j|v, w; R], u; R] \\ &\quad + [i|\,[j|\, w, u; R], v; R] + [j|\,[i|\, u, v; R], w; R] + \\ &\quad + [j|\,[i|v, w; R], u; R] + [j|\,[i|\, w, u; R], v; R] = 0. \end{aligned}$$

Proof. (a) follows directly from the definition by expansion. (b) follows from the antisymmetry of $\vec{\boldsymbol{I}}$ in a manner analogous to the proof of part (b) of Theorem 4.3. (c) through (h) follow directly from (a) upon noting that p and q are independent in I. (i): Let

$$\begin{aligned} (i, j|\, v, w; u) = u_{,p_{\alpha j} p_{\beta i}} v_{,q_\alpha} w_{,q_\beta} - u_{,q_\alpha p_{\beta i}} v_{,p_{\alpha j}} w_{,q_\beta} - \\ - u_{,p_{\alpha j} q_\beta} v_{,q_\alpha} w_{,p_{\beta i}} + u_{,q_\alpha q_\beta} v_{,p_{\alpha j}} w_{,p_{\beta i}}; \end{aligned}$$

[17] POISSON, S. D.: J. de l'École Polytech. VIII (1809).

then

$$[i|[j|u,v;R],w;R] = (i,j|v,w;u) - (i,j|u,w;v).$$

Noting that $(i,j|v,w;u) = (j,i|w,v;u)$, the result follows. Q.E.D.

Part (i) of this theorem is JACOBI's identity generalized to the case where $n > 1$.

The question concerning the nature of the image of the Poisson bracket vectrix under an element of the generalized symplectic group now arises. Under a map T of the collection $\mathscr{T}(T:\{R\} \to \{\tilde{R}\})$ we have

$$u_{,\{R\}} = u_{,\{\tilde{R}\}} \{\tilde{R}\}_{,\{R\}} = u_{,\{\tilde{R}\}} \boldsymbol{M}^{-1}$$

since $\boldsymbol{M}^{-1} = \{\tilde{R}\}_{,\{R\}}$. Thus, from the definition of the Poisson bracket vectrix, we find that

$$\vec{P}(u,v;R) = u_{,\{\tilde{R}\}} \boldsymbol{M}^{-1} \vec{\boldsymbol{I}} \boldsymbol{M}^{-1\prime} \{v_{,\tilde{R}}\}.$$

With respect to $\{\tilde{R}\}$ we have $\vec{P}(u,v;\tilde{R}) = u_{,\{\tilde{R}\}} \vec{\boldsymbol{I}} \{v_{,\tilde{R}}\}$, so that $\vec{P}(u,v;R)$ will change by a constant multiple ψ under T if and only if

$$u_{,\{\tilde{R}\}} \boldsymbol{M}^{-1} \vec{\boldsymbol{I}} \boldsymbol{M}^{-1\prime} \{v_{,\tilde{R}}\} = \psi\, u_{,\{\tilde{R}\}} \vec{\boldsymbol{I}} \{v_{,\tilde{R}}\}$$

is an identity in I. The most general condition under which the above identity is satisfied is given by

$$\boldsymbol{M}^{-1\prime} \vec{\boldsymbol{I}} \boldsymbol{M}^{-1\prime} = \psi \vec{\boldsymbol{I}} + \vec{\boldsymbol{\beta}},$$

where $\vec{\boldsymbol{\beta}}$ is such that $u_{,\{\tilde{R}\}} \vec{\boldsymbol{\beta}} \{v_{,\tilde{R}}\} = \vec{\boldsymbol{O}}$. Now, if this condition is to be satisfied for all choices of u and v, it must also be satisfied for u and v elements of $\{\tilde{R}\}$. Thus, it is required that $\{\tilde{R}\}_{,\{\tilde{R}\}} \vec{\boldsymbol{\beta}} \{\tilde{R}\}_{,\{\tilde{R}\}} = \vec{\boldsymbol{O}}$, from which we may conclude that $\vec{\boldsymbol{\beta}}$ must be a zero vectrix, and hence $\boldsymbol{M}^{-1} \vec{\boldsymbol{I}} \boldsymbol{M}^{-1\prime} = \psi \vec{\boldsymbol{I}}$ is a necessary and sufficient condition that the Poisson bracket vectrix changes by the multiplicative factor ψ under T. Premultiplying by $\boldsymbol{M}$ and postmultyplying by $\boldsymbol{M}'$ gives $\boldsymbol{M}\vec{\boldsymbol{I}}\boldsymbol{M}' = \frac{1}{\psi}\vec{\boldsymbol{I}}$ as a necessary and sufficient condition that the Poisson bracket vectrix change by a constant multiple only under T. If T be an element of the generalized symplectic group, then $\boldsymbol{M}$ must satisfy, in addition to the above equation, the equation $\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} = \frac{1}{\mu}\vec{\boldsymbol{I}}$ by definition. Hence, the Poisson bracket vectrix changes by a multiplicative constant under the generalized symplectic group if and only if the associated Jacobian matrices of the elements of the group satisfy the simultaneous equations $\boldsymbol{M}\vec{\boldsymbol{I}}\boldsymbol{M}' = \frac{1}{\psi}\vec{\boldsymbol{I}}$, $\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} = \frac{1}{\mu}\vec{\boldsymbol{I}}$. It is immediate that $\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} = \frac{1}{\mu}\vec{\boldsymbol{I}}$ implies $\boldsymbol{M}\vec{\boldsymbol{I}}\boldsymbol{M}' = \frac{1}{\psi}\vec{\boldsymbol{I}}$ if $\boldsymbol{M}$ is either symmetric or anti-symmetric, upon setting $\psi = \mu$, and if $\boldsymbol{M}$ is orthogonal upon setting $\psi = 1/\mu$. Setting

$$\boldsymbol{M} = \begin{pmatrix} \xi^{1} & 0 & 0 \\ \lambda_1 \xi^{-1} & \xi & 0 \\ \lambda_2 \xi^{1} & 0 & \xi \end{pmatrix},$$

it is evident that such an $\boldsymbol{M}$ satisfies the equation $\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M}=\vec{\boldsymbol{I}}$, so that it is an element of the generalized symplectic group. On the other hand

$$\boldsymbol{M}\vec{\boldsymbol{I}}\boldsymbol{M}'=\begin{pmatrix} 0 & \vec{e}_1 & \vec{e}_2 \\ -\vec{e}_1 & 0 & -\vec{e}_1\lambda_2+\vec{e}_2\lambda_1 \\ -\vec{e}_2 & \vec{e}_1\lambda_2-\vec{e}_2\lambda_1 & 0 \end{pmatrix}+\frac{1}{\psi}\vec{\boldsymbol{I}},$$

and hence, in this case, $\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M}=\frac{1}{\mu}\vec{\boldsymbol{I}}$ does not imply $\boldsymbol{M}\vec{\boldsymbol{I}}\boldsymbol{M}'=\frac{1}{\psi}\vec{\boldsymbol{I}}$, so that the Poisson bracket vector does not admit the property of changing by a multiplicative constant factor under every element of the generalized symplectic group.

In the case $n=1$, $\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M}=\frac{1}{\mu}\vec{\boldsymbol{I}}$ implies $\boldsymbol{M}\vec{\boldsymbol{I}}\boldsymbol{M}'=\frac{1}{\mu}\vec{\boldsymbol{I}}$. Specifically, $\vec{\boldsymbol{I}}_l^{-1}=\vec{\boldsymbol{I}}_r^{-1}=-\vec{\boldsymbol{I}}$ as seen from equation (2.6), for $n=1$, and hence $\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M}=\frac{1}{\mu}\vec{\boldsymbol{I}}$ implies $\boldsymbol{M}'=\frac{1}{\mu}\vec{\boldsymbol{I}}\boldsymbol{M}^{-1}\bullet\vec{\boldsymbol{I}}_r^{-1}=-\frac{1}{\mu}\vec{\boldsymbol{I}}\boldsymbol{M}^{-1}\bullet\vec{\boldsymbol{I}}=\frac{1}{\mu}\vec{\boldsymbol{I}}_l^{-1}\bullet\boldsymbol{M}^{-1}\vec{\boldsymbol{I}}$, whence $\boldsymbol{M}\vec{\boldsymbol{I}}\boldsymbol{M}'=\frac{1}{\mu}\vec{\boldsymbol{I}}$ since there is no $\vec{\boldsymbol{C}}$ such that $\vec{\boldsymbol{C}}\bullet\vec{\boldsymbol{I}}=\vec{\boldsymbol{O}}$ other than $\vec{\boldsymbol{C}}=\vec{\boldsymbol{O}}$ in this case.

We have thus proved

Theorem 4.12. *For n greater than one, Poisson bracket vectrices do not admit the property of changing by a constant multiplicative factor under all elements of the generalized symplectic group, although this property does result if the associated Jacobian matrix is either symmetric, antisymmetric, or orthogonal.*

Theorem 4.12 constitutes the first major difference between the properties of the classic symplectic group and the generalized symplectic group. The difference, namely the lack of invariance of the Poisson bracket vectrix under the basic symplectic group as compared with the invariance of the Poisson bracket for the classic symplectic group, is of considerable moment in that it is the Poisson bracket vectrix or its functional equivalent which is used to form the commutation relations for quantum field theory.

Left and Right-Hand Bracket Vectrices

To alleviate the situation pointed out by Theorem 4.12, we inquire as to whether it is possible to construct bracket vectrices which will reduce to the Poisson bracket for the case $n=1$ and which, in addition, admit the property of changing by a constant multiplicative factor under any element of the generalized symplectic group.

In the classic case ($n=1$), the Poisson brackets are the reciprocals of the Lagrange brackets. We thus investigate the problem of constructing bracket vectrices which are the reciprocals of the Lagrange bracket vectrices for the case $n>1$. Let $\{u\}$ be an $N(n+1)$-element column matrix of functions of $\{R\}$ such that the matrix $\boldsymbol{V}=\{u\}_{,\{R\}}$ is nonsingular. From the definition of $\vec{L}(u,v)$ we have $\vec{L}(\{u\},\{u\})=(\{R\}_{,\{u\}})'\vec{\boldsymbol{I}}\{R\}_{,\{u\}}$. Set $\mathscr{L}(\{u\},\{u\};R)=\{u\}_{,\{R\}}\vec{\boldsymbol{A}}(\{u\}_{,\{R\}})'$ where $\vec{\boldsymbol{A}}$ is to be determined by the requirement

$$\vec{\mathscr{L}}(\{u\},\{u\};R)\bullet\vec{L}(\{u\},\{u\})=\boldsymbol{E}^*. \tag{4.20}$$

From the above forms we have $\vec{\mathscr{L}}(\{u\},\{u\};R)\bullet\vec{L}(\{u\},\{u\})=\boldsymbol{V}\vec{\boldsymbol{A}}\boldsymbol{V}'\bullet\boldsymbol{V}^{-1'}\vec{\boldsymbol{I}}\boldsymbol{V}^{-1}=\boldsymbol{V}\vec{\boldsymbol{A}}\bullet\vec{\boldsymbol{I}}=\boldsymbol{E}^*$ in order for equation (4.20) to be satisfied. Multiplying the last equation from the left by $\boldsymbol{V}$ gives $\boldsymbol{V}(\vec{\boldsymbol{A}}\bullet\vec{\boldsymbol{I}})=\boldsymbol{V}$, from which we may conclude that $\vec{\boldsymbol{A}}$ must be such as to satisfy $\vec{\boldsymbol{A}}\bullet\vec{\boldsymbol{I}}=\boldsymbol{E}^*$ in order for equation (4.20) to hold.

We now inquire into what form $\vec{\mathscr{L}}$ must have in order for it to change by a constant multiplicative factor under an element of the generalized symplectic group. From the definition of $\vec{\mathscr{L}}$ we have

$$\vec{\mathscr{L}}(\{u\},\{u\};R)=\{u\}_{,\{R\}}\vec{\boldsymbol{A}}(\{u\}_{,\{R\}})'.$$

Under a canonical map of multiplier μ we obtain $\{u\}_{,\{R\}}=\{u\}_{,\{\tilde{R}\}}\boldsymbol{M}^{-1}$, and hence

$$\vec{\mathscr{L}}(\{u\},\{u\};R)=\{u\}_{,\{\tilde{R}\}}\boldsymbol{M}^{-1}\vec{\boldsymbol{A}}\boldsymbol{M}^{-1'}(\{u\}_{,\{\tilde{R}\}})'.$$

If we set $\vec{\mathscr{L}}(\{u\},\{u\};\tilde{R})=\{u\}_{,\{\tilde{R}\}}\tilde{\vec{A}}(\{u\}_{,\{\tilde{R}\}})'$, then $\mathscr{L}$ will change by a constant multiplicative factor η under the canonical map considered if and only if

$$\begin{aligned}&\boldsymbol{M}^{-1}\vec{\boldsymbol{A}}\boldsymbol{M}^{-1'}=\eta\,\tilde{\vec{\boldsymbol{A}}}+\vec{\boldsymbol{\beta}},\\&\{u\}_{,\{\tilde{R}\}}\vec{\boldsymbol{\beta}}(\{u\}_{,\{\tilde{R}\}})'=\boldsymbol{O}\end{aligned}\tag{4.21}$$

is satisfied. If we require equations (4.21) to hold independently of the choice of $\{u\}$ (that is, that $\vec{\mathscr{L}}(u,u;R)$ changes by a constant multiplicative factor under the map considered for all $\{u\}$ satisfying the condition that $\{u\}_{,\{R\}}$ is nonsingular), then it must hold for $\{u\}=\{\tilde{R}\}$, from which we conclude that $\vec{\boldsymbol{\beta}}$ must be a zero vectrix.

The map considered was assumed to be canonical, so that $\boldsymbol{M}$ satisfies $\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M}=\frac{1}{\mu}\vec{\boldsymbol{I}}$. Thus, $\vec{\boldsymbol{I}}=\frac{1}{\mu}\boldsymbol{M}^{-1'}\vec{\boldsymbol{I}}\boldsymbol{M}^{-1}$ from which we obtain $\boldsymbol{E}^*=\frac{1}{\mu}\vec{\boldsymbol{A}}\bullet\boldsymbol{M}^{-1'}\vec{\boldsymbol{I}}\boldsymbol{M}^{-1}$ upon noting that $\vec{\boldsymbol{A}}\bullet\vec{\boldsymbol{I}}=\boldsymbol{E}^*$. This result may be written in the more useful form $\boldsymbol{M}=\frac{1}{\mu}\vec{\boldsymbol{A}}\boldsymbol{M}^{-1'}\bullet\vec{\boldsymbol{I}}$. From equation (4.21) we obtain by direct matrix operations $\vec{\boldsymbol{A}}\boldsymbol{M}^{-1'}=\eta\boldsymbol{M}\tilde{\vec{\boldsymbol{A}}}$, which, upon substitution into the above equation, gives $\boldsymbol{M}=\frac{\eta}{\mu}\boldsymbol{M}\tilde{\vec{\boldsymbol{A}}}\bullet\vec{\boldsymbol{I}}$. If we require $\tilde{\vec{\boldsymbol{A}}}$ to be such that it also satisfies $\tilde{\vec{\boldsymbol{A}}}\bullet\vec{\boldsymbol{I}}=\boldsymbol{E}^*$ (that is, $\vec{\mathscr{L}}(u,u;\tilde{R})$ is the left inverse of $\vec{L}(u,u;\tilde{R})$ under the map considered), then we have $\boldsymbol{M}=\frac{\eta}{\mu}\boldsymbol{M}$ so that $\eta=\mu$.

Set

$$\vec{\mathscr{R}}(\{u\},\{u\};R)=-\vec{\mathscr{L}}'(\{u\},\{u\};R),$$

so that

$$\vec{\mathscr{R}}(\{u\},\{u\};R)=-\{u\}_{,\{R\}}\vec{\boldsymbol{A}}'(\{u\}_{,\{R\}})'.$$

Since $\vec{\boldsymbol{A}}\bullet\vec{\boldsymbol{I}}=\boldsymbol{E}^*$, it follows that $\vec{\boldsymbol{I}}\bullet(-\vec{\boldsymbol{A}}')=\boldsymbol{E}^*$ owing to the fact that $\vec{\boldsymbol{I}}'=-\vec{\boldsymbol{I}}$ and $\boldsymbol{E}^{*\prime}=\boldsymbol{E}^*$. Thus, we see that

$$\vec{L}(\{u\},\{u\})\bullet\vec{\mathscr{R}}(\{u\},\{u\};R)=\boldsymbol{E}^*.$$

Summarizing, we have the following

*Theorem 4.13. *Let $\{u(R, \boldsymbol{x})\}$ be an $N(n+1)$-element column matrix of functions of class C^1 such that the matrix $\{u\}_{,\{R\}}$ is nonsingular. Define the* ***right and left-hand bracket vectrices by***

$$\vec{\mathscr{R}}(w, v; R) = -v_{,\{R\}} \vec{\boldsymbol{A}}' \{w_{,R}\}, \tag{4.22}$$

$$\vec{\mathscr{L}}(w, v; R) = v_{,\{R\}} \vec{\boldsymbol{A}} \{w_{,R}\} \tag{4.23}$$

where $\vec{\boldsymbol{A}}$ satisfies

$$\vec{\boldsymbol{A}} \bullet \vec{\boldsymbol{I}} = \boldsymbol{E}^*; \tag{4.24}$$

then

$$\begin{aligned} \vec{\mathscr{L}}(\{u\}, \{u\}; R) \bullet \vec{L}(\{u\}, \{u\}) &= \boldsymbol{E}^*, \\ \vec{L}(\{u\}, \{u\}) \bullet \vec{\mathscr{R}}(\{u\}, \{u\}; R) &= \boldsymbol{E}^*. \end{aligned} \tag{4.25}$$

In addition, the images of $\vec{\mathscr{L}}(v, w; R)$ and $\vec{\mathscr{R}}(v, w; R)$ are equal to $\eta\, \vec{\mathscr{L}}(v, w; \tilde{R})$ and $\eta\, \vec{\mathscr{R}}(v, w; \tilde{R})$ respectively, and equations (4.20) *are invariant under a canonical map of multiplier μ if and only if*

(i) $$\eta = \mu,$$

(ii) $$\boldsymbol{M}^{-1} \vec{\boldsymbol{A}} \boldsymbol{M}^{-1\prime} = \eta\, \tilde{\vec{\boldsymbol{A}}} \tag{4.26}$$

where $\tilde{\vec{\boldsymbol{A}}} \bullet \vec{\boldsymbol{I}} = \boldsymbol{E}^*$ and where

$$\vec{\mathscr{L}}(\{u\}, \{u\}; \tilde{R}) = \{u\}_{,\{\tilde{R}\}} \tilde{\vec{\boldsymbol{A}}}(\{u\}_{,\{\tilde{R}\}}). \tag{4.27}$$

We first show that the collection of $\vec{\boldsymbol{A}}$ which satisfies the conditions of the last theorem is nonvacuous. From the properties of $\vec{\boldsymbol{I}}_l^{-1}$ we have $\vec{\boldsymbol{I}}_l^{-1} \bullet \vec{\boldsymbol{I}} = \boldsymbol{E}^*$, and hence by setting $\vec{\boldsymbol{A}} = \vec{\boldsymbol{I}}_l^{-1} + \vec{\boldsymbol{C}}$, where $\vec{\boldsymbol{C}}$ is any vectrix such that $\vec{\boldsymbol{C}} \bullet \boldsymbol{I} = \boldsymbol{O}$, we satisfy the condition stated by equations (4.24).

From the definition of a symplectic matrix $\left(\boldsymbol{M}' \vec{\boldsymbol{I}} \boldsymbol{M} = \frac{1}{\mu} \vec{\boldsymbol{I}}\right)$ we have $\vec{\boldsymbol{A}}^* \bullet \boldsymbol{M}' \vec{\boldsymbol{I}} \boldsymbol{M} = \frac{1}{\mu} \boldsymbol{E}^*$ by equation (4.24). Hence, $\vec{\boldsymbol{A}}^* \bullet \boldsymbol{M}' \vec{\boldsymbol{I}} = \frac{1}{\mu} \boldsymbol{M}^{-1}$ so that $\boldsymbol{M} \vec{\boldsymbol{A}}^* \boldsymbol{M}' \bullet \vec{\boldsymbol{I}} = \frac{1}{\mu} \boldsymbol{E}^* = \frac{1}{\mu} \vec{\boldsymbol{A}}^* \bullet \vec{\boldsymbol{I}}$ since $\vec{\boldsymbol{A}}^* \bullet \vec{\boldsymbol{I}} = \boldsymbol{E}^*$. Thus we obtain

$$\left(\boldsymbol{M} \vec{\boldsymbol{A}}^* \boldsymbol{M}' - \frac{1}{\mu} \vec{\boldsymbol{A}}^*\right) \bullet \vec{\boldsymbol{I}} = \boldsymbol{O},$$

which by Theorem 1.2 gives

$$\boldsymbol{M} \vec{\boldsymbol{A}}^* \boldsymbol{M}' = \frac{1}{\mu} (\vec{\boldsymbol{A}}^* + \vec{\boldsymbol{C}}),$$

where $\vec{\boldsymbol{C}}$ is arbitrary to within the condition $\vec{\boldsymbol{C}} \bullet \vec{\boldsymbol{I}} = \boldsymbol{O}$. Hence, $\boldsymbol{M}' \boldsymbol{I} \boldsymbol{M} = \frac{1}{\mu} \vec{\boldsymbol{I}}$ implies $\boldsymbol{M} \vec{\boldsymbol{A}}^* \boldsymbol{M}' = \frac{1}{\mu} (\vec{\boldsymbol{A}}^* + \vec{\boldsymbol{C}})$. Identifying $\vec{\boldsymbol{A}}^*$ with $\tilde{\vec{\boldsymbol{A}}}$ and $\vec{\boldsymbol{A}}^* + \vec{\boldsymbol{C}}$ with $\vec{\boldsymbol{A}}$, which is permissible since $(\vec{\boldsymbol{A}}^* + \vec{\boldsymbol{C}}) \bullet \vec{\boldsymbol{I}} = \vec{\boldsymbol{A}}^* \bullet \vec{\boldsymbol{I}} = \boldsymbol{E}^*$, we see that equations (4.24) and (4.26) are satisfied. We have thus proved

*Theorem 4.14. *Let $\boldsymbol{M}$ be a generalized symplectic matrix of multiplier μ, and let $\vec{\boldsymbol{A}}$ be any vectrix such that*

$$\vec{\boldsymbol{A}} \bullet \vec{\boldsymbol{I}} = \boldsymbol{E}^*; \tag{4.28}$$

then there exists a vectrix $\vec{C}$, for every $\boldsymbol{M}$ such that

$$\vec{C} \bullet \vec{I} = O \tag{4.29}$$

for which $\boldsymbol{M}'\vec{I}\boldsymbol{M} = \frac{1}{\mu}\vec{I}$ implies

$$\boldsymbol{M}^{-1}\vec{A}\boldsymbol{M}^{-1\prime} = \mu\,\tilde{\tilde{\vec{A}}} \tag{4.30}$$

where

$$\tilde{\tilde{\vec{A}}} = \vec{A} + \vec{C}, \tag{4.31}$$

and conversely. In particular, we may take $\vec{A} = \vec{I}_l^{-1}$.

***Corollary.** *The right and left-hand bracket vectrices change only by a multiplicative factor μ under every element of the generalized symplectic group of multiplier μ.*

Proof. The proof follows directly from Theorems 4.13 and 4.14 provided we show that equation (4.30) holds for products of elements of the generalized symplectic group. Let

$$\boldsymbol{M}_1\vec{I}_l^{-1}\boldsymbol{M}_1' = \frac{1}{\mu_1}(\vec{I}_l^{-1} + \vec{C}_1),$$

$$\boldsymbol{M}_2\vec{I}_l^{-1}\boldsymbol{M}_2' = \frac{1}{\mu_2}(\vec{I}_l^{-1} + \vec{C}_2);$$

then

$$\begin{aligned}\boldsymbol{M}_2\boldsymbol{M}_1\vec{I}_l^{-1}\boldsymbol{M}_1'\boldsymbol{M}_2' &= \frac{1}{\mu_1}\boldsymbol{M}_2(\vec{I}_l^{-1} + \vec{C}_1)\boldsymbol{M}_2' \\ &= \frac{1}{\mu_1}\boldsymbol{M}_2\vec{I}_l^{-1}\boldsymbol{M}_2' + \frac{1}{\mu_1}\boldsymbol{M}_2\vec{C}_1\boldsymbol{M}_2' \\ &= \frac{1}{\mu_1\mu_2}(\vec{I}_l^{-1} + \vec{C}_2) + \frac{1}{\mu_1}\boldsymbol{M}_2\vec{C}_1\boldsymbol{M}_2'.\end{aligned}$$

Therefore, we must show that $\tilde{\vec{C}} = \boldsymbol{M}_2\vec{C}_1\boldsymbol{M}_2'$ is such that $\tilde{\vec{C}} \bullet \vec{I} = O$. Now $\tilde{\vec{C}} \bullet \vec{I} = \boldsymbol{M}_2\vec{C}_1 \bullet \boldsymbol{M}_2'\vec{I}$, but $\boldsymbol{M}_2'\vec{I} = \frac{1}{\mu_2}\vec{I}\boldsymbol{M}_2^{-1}$, since T is canonical, and hence $\tilde{\vec{C}} = \frac{1}{\mu_2}\boldsymbol{M}_2\vec{C}_1 \bullet \vec{I}\boldsymbol{M}_2^{-1} = O$ since $\boldsymbol{M}$ is nonsingular. Therefore, setting $\boldsymbol{M} = \boldsymbol{M}_2\boldsymbol{M}_1$, we have

$$\boldsymbol{M}\vec{I}_l^{-1}\boldsymbol{M}' = \frac{1}{\mu_1\mu_2}(\vec{I}_l^{-1} + \vec{C})$$

where

$$\vec{C} = \vec{C}_2 + \mu_2\boldsymbol{M}_2\vec{C}_1\boldsymbol{M}_2'. \qquad \text{Q.E.D.}$$

These equations constitute the combinational laws for the $\vec{C}$ under $\boldsymbol{M}_2\boldsymbol{M}_1$.

In the classic case ($n = 1$) the equation $\vec{A} \bullet \vec{I} = E^*$ implies that $\vec{A} = -\vec{I}$ since in this case $\vec{I} = \vec{e}_1\begin{pmatrix} O & E \\ -E & O\end{pmatrix}$ (that is, $\vec{I}_r^{-1} = \vec{I}_l^{-1} = -\vec{I}$), and there is no vectrix $\vec{C}$ other than the zero vectrix such that $\vec{C} \bullet \vec{I} = O$. Thus, the right and left-hand bracket vectrices reduce to the Poisson bracket vectrix for $n = 1$, as is evident by using the above results in comparing equations (4.19), (4.22), and (4.23). We have thus obtained the desired generalization of the Poisson bracket vectrix for the case $n > 1$.

It is of interest to note that for the matrix

$$\boldsymbol{M} = \begin{pmatrix} \xi^{-1} & 0 & 0 \\ \lambda_1 \xi^{-1} & \xi & 0 \\ \lambda_2 \xi^{-1} & 0 & \xi \end{pmatrix},$$

which is canonical with multiplier unity, we have

$$\boldsymbol{M} \vec{\boldsymbol{I}}_l^{-1} \boldsymbol{M}' = \begin{pmatrix} 0 & -\vec{e}_1/2 & -\vec{e}_2/2 \\ \vec{e}_1 & \vec{e}_1 \lambda_1/2 & \vec{e}_1 \lambda_2 - \vec{e}_2 \lambda_1/2 \\ \vec{e}_2 & -\vec{e}_1 \lambda_1/2 + \vec{e}_2 \lambda_1 & \vec{e}_2 \lambda_2/2 \end{pmatrix}$$

so that

$$\vec{\boldsymbol{C}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \vec{e}_1 \lambda_1/2 & \vec{e}_1 \lambda_2 - \vec{e}_2 \lambda_1/2 \\ 0 & -\vec{e}_1 \lambda_2/2 + \vec{e}_2 \lambda_1 & \vec{e}_2 \lambda_2/2 \end{pmatrix},$$

for which one may easily verify that $\vec{\boldsymbol{C}} \bullet \vec{\boldsymbol{I}} = \boldsymbol{O}$.

The properties of the right and left-hand bracket vectrices are as follows:

Theorem 4.15. *Let $u(R, \boldsymbol{x})$, $v(R, \boldsymbol{x})$ be arbitrary functions of class C^1; then*

*(a) $$\vec{\mathscr{R}}(u, v; R) = -\vec{\mathscr{L}}'(v, u; R). \tag{4.32}$$

With $\vec{\boldsymbol{A}} = \vec{\boldsymbol{I}}_l^{-1}$, we have

(b) $$\vec{\mathscr{L}}(u, v; R) = \langle i | u, v; R \rangle \vec{e}_i \tag{4.33}$$

where

$$\langle i | u, v; R \rangle = u_{,\mathrm{q}_\alpha} v_{,p_{\alpha i}} - \frac{1}{n} u_{,p_{\alpha i}} v_{,\mathrm{q}_\alpha},$$

(c) $$\vec{\mathscr{L}}(p_{\beta j}, p_{\gamma k}; R) \equiv 0, \tag{4.34}$$

(d) $$\vec{\mathscr{L}}(\mathrm{q}_\beta, \mathrm{q}_\alpha; R) \equiv 0, \tag{4.35}$$

(e) $$\begin{aligned} \vec{\mathscr{L}}(\mathrm{q}_\beta, p_{\gamma j}; R) &= \delta_{\beta\gamma} \vec{e}_j, \\ \vec{\mathscr{L}}(p_{\gamma j}, \mathrm{q}_\beta; R) &= \frac{-1}{n} \delta_{\beta\gamma} \vec{e}_j, \end{aligned} \tag{4.36}$$

*(f) $$\vec{\mathscr{L}}\big(f(u_1, \ldots, u_s), v; R\big) = \sum_{\xi=1}^{s} f_{,u_\xi} \vec{\mathscr{L}}(u_\xi, v; R), \tag{4.37}$$

(g) $$\vec{\mathscr{L}}(u, u; R) = \frac{n-1}{n} u_{,\mathrm{q}_\alpha} u_{,p_{\alpha i}} \vec{e}_i, \tag{4.38}$$

(h) $$\begin{aligned} \vec{\mathscr{L}}(u, v; R) = \big[&\vec{\mathscr{L}}(\mathrm{q}_\beta, u; R) \bullet \vec{\mathscr{L}}(p_{\beta j}, v; R) - \\ &- n \vec{\mathscr{L}}(\mathrm{q}_\beta, v; R) \bullet \vec{\mathscr{L}}(p_{\beta j}, u; R)\big] \vec{e}_j. \end{aligned} \tag{4.39}$$

Proof. (a) From the definition

$$\vec{\mathscr{R}}'(u, v; R) = \big(-v_{,\{R\}} \vec{\boldsymbol{A}}' \{u_{,R}\}\big)' = -u_{,\{R\}} \vec{\boldsymbol{A}} \{v_{,R}\},$$

which gives $\mathscr{R}'(u, v; R) = -\vec{\mathscr{L}}(v, u; R)$. (b) follows directly from equation (4.23) upon expansion. (c), (d), and (e) follow directly from (b) upon noting that p and q are independent in I. (f) and (g) are immediate from the definition of $\vec{\mathscr{L}}$. Q.E.D.

Equations (4.33) admit rather interesting results. One obtains, by use of these equations and by noting that (p, q) are independent in I,

$$\vec{\mathscr{L}}(p_{\alpha j}, H; R) = \frac{-1}{n} H_{,q_\beta} \vec{e}_j, \tag{4.40}$$

$$\vec{\mathscr{L}}(q_\alpha, H; R) = H_{,p_{\beta i}} \vec{e}_i. \tag{4.41}$$

Hence, by equations (2.12), if (p, q) form a Hamiltonian system base H, then (p, q) must satisfy the equations

$$\vec{\nabla} q_\alpha = \vec{\mathscr{L}}(q_\alpha, H; R), \tag{4.42}$$

$$\vec{\nabla} \bullet \vec{p}_\alpha = \vec{\mathscr{L}}(\bullet \vec{p}, H; R), \tag{4.43}$$

where

$$\begin{gathered} \vec{\mathscr{L}}(\bullet \vec{a}, b; R) = b_{,\{R\}} \vec{A} \bullet \{\vec{a}_{,R}\} \\ \vec{p}_\alpha = \vec{e}_i p_{\alpha i}. \end{gathered} \tag{4.44}$$

A Subsidiary Result

***Theorem 4.16.** *Let $\{dR\}$ and $\{\Delta R\}$ represent independent infinitesimal changes in the elements of $\{R\} \in I$. A map T of the collection $\mathscr{T}$ is a canonical map of multiplier μ if and only if the vectrix equation*

$$\vec{\mathscr{D}}(d, \Delta; R) = \frac{1}{\mu} \vec{\mathscr{D}}(d, \Delta; \tilde{R}) \tag{4.45}$$

is satisfied for all $\{R\} \in I$, where

$$\begin{aligned} \vec{\mathscr{D}}(d, \Delta; R) &= \{dR\}' \vec{\boldsymbol{I}} \{\Delta R\}, \\ \vec{\mathscr{D}}(d, \Delta; \tilde{R}) &= \{d\tilde{R}\}' \vec{\boldsymbol{I}} \{\Delta \tilde{R}\}. \end{aligned} \tag{4.46}$$

Proof. Under T

$$\{dR\} = \{R\}_{,\{\tilde{R}\}} \{d\tilde{R}\} = \boldsymbol{M}\{d\tilde{R}\},$$

$$\{\Delta R\} = \boldsymbol{M}\{\Delta \tilde{R}\},$$

and hence

$$\vec{\mathscr{D}}(d, \Delta; R) = \{d\tilde{R}\}' \boldsymbol{M}' \vec{\boldsymbol{I}} \boldsymbol{M} \{\Delta \tilde{R}\},$$

from which the result follows since

$$\boldsymbol{M}' \vec{\boldsymbol{I}} \boldsymbol{M} = \frac{1}{\mu} \vec{\boldsymbol{I}}$$

is both a necessary and a sufficient condition that T is canonical. Q.E.D.

This theorem is analogous to the characterization of the classic symplectic group by means of Pfaffians, except that in the case $n > 1$ the antisymmetric character is reflected in the antisymmetric character of the structure vectrix for the system of equations considered.

Section V. Partition Spaces and Partition Maps

Canonical Maps and Partition Spaces

In this section we consider those maps of the collection $\mathscr{T}$ which are such that $p_{\alpha i}$ depends only on Q_β and $P_{\beta i}$ for fixed i, so that $\partial p_{\alpha i}/\partial P_{\beta j}$ is zero for all $i \neq j$.

Definition. By the n **partition spaces** of the generalized phase space I, denoted by $\mathscr{P}^i(p, \mathrm{q})$, are meant the n $2N$-dimensional subspaces of I with coordinates $(\mathrm{q}_\alpha, p_{\alpha i})$ for fixed $i = 1, \ldots, n$. It is to be noted that the intersection $\mathscr{P}^i \cap \mathscr{P}^j$ for all $i \neq j$ is just the subspace of I with coordinates (q_α) and that the union $\bigcup_i \mathscr{P}^i$ over all i is the space I.

Definition. By a **canonical partition map** is meant a canonical map which maps the partition space $\mathscr{P}^i(p, \mathrm{q})$ of I for each $i = 1, \ldots, n$ into the corresponding **partition space** $\mathscr{P}^i(P, Q)$.

In Section II we introduced the partition of the associated Jacobian matrix of a canonical map by the equation

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}.$$

Thus, under a canonical partition map, the $nN \times nN$ matrix $\boldsymbol{D}$ has only diagonal $N \times N$ matrices which are non-zero. It is to be noted that the collection of all canonical partition maps does not form a group, since the product of the associated Jacobian matrices of two canonical partition maps does not result in a partition map in general (that is, since

$$\begin{pmatrix} \boldsymbol{A}_2 & \boldsymbol{B}_2 \\ \boldsymbol{C}_2 & \boldsymbol{D}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{B}_1 \\ \boldsymbol{C}_1 & \boldsymbol{D}_1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}_2 \boldsymbol{A}_1 + \boldsymbol{B}_2 \boldsymbol{C}_1 & \boldsymbol{A}_2 \boldsymbol{B}_1 + \boldsymbol{B}_2 \boldsymbol{D}_1 \\ \boldsymbol{C}_2 \boldsymbol{A}_1 + \boldsymbol{D}_2 \boldsymbol{C}_1 & \boldsymbol{C}_2 \boldsymbol{B}_1 + \boldsymbol{D}_2 \boldsymbol{D}_1 \end{pmatrix},$$

the product of two canonical partition maps will be a partition map if and only if either $\boldsymbol{C}_2$ or $\boldsymbol{B}_1$ is a matrix all of whose elements are zero).

Theorem 5.1. *Let s_2^i be a smooth two-dimensional subspace of $\mathscr{P}^i(p, \mathrm{q})$, and set*

$$\mathscr{I}^i = \int_{s_2^i} d p_{\alpha i}\, d \mathrm{q}_\alpha. \tag{5.1}$$

Denote the image of $\mathscr{I}^i$ under a canonical partition map of multiplier μ by $\widetilde{\mathscr{I}^i}$; then

$$\mathscr{I}^i = \frac{1}{\mu} \widetilde{\mathscr{I}^i}. \tag{5.2}$$

Proof. Since s_2^i is a smooth two-dimensional surface, we may characterize it by a continuous two parameter family u, v and thus

$$\begin{aligned} \mathscr{I}^i &= \int_{s_2^i} \frac{\partial(p_{\alpha i}, \mathrm{q}_\alpha)}{\partial(u, v)}\, du\, dv \\ &= \int_{s_2^i} (p_{\alpha i, u} \mathrm{q}_{\alpha, v} - p_{\alpha i, v} \mathrm{q}_{\alpha, u})\, du\, dv \\ &= \int_{s_2^i} \{i \,|\, u, v; p\, \mathrm{q}\}\, du\, dv, \end{aligned}$$

so that $\mathscr{I}^i$ is just the integral of the negative of the i^{th} component of the Lagrange bracket vectrix. Hence, by Theorem 4.3 (h)

$$\{i \mid u, v; p\, \mathrm{q}\} = \frac{1}{\mu} \{i \mid u, v; P\, Q\},$$

and so

$$\begin{aligned}\mathscr{I}^i &= \frac{1}{\mu} \int_{s_2^i} \{i \mid u, v; P\, Q\}\, du\, dv \\ &= \frac{1}{\mu} \int_{s_2^i} \frac{\partial(P_{\alpha i}, Q_\alpha)}{\partial(u, v)}\, du\, dv = \frac{1}{\mu} \tilde{\mathscr{I}}^i.\end{aligned}$$

Q.E.D.

Let the numerosity of the α index set, $\mathscr{A}$, be N and consider the following N-vectors in the generalized phase space I:

(a) the N-vector

$$\mathrm{q} = \{\mathrm{q}_\alpha\};$$

(b) the n N-vectors

$$p_i = \{p_{\alpha i}\}.$$

From these N-vectors we may construct n $2N$-vectors $\{z_i\}$ whose components are defined by

$$z_{i\alpha} = \mathrm{q}_\alpha; \qquad z_{i\alpha+N} = p_{\alpha i}.$$

It is to be noted that $\{z_i\}$ spans the partition space $\mathscr{P}^i(p, \mathrm{q})$.

Let T be a canonical partition map so that T maps z_i to Z_i. (That is, $\{z_i\} = \{z_i(Z, \boldsymbol{x})\}$.) Set $(\Gamma_{ij}) = \{z_i\}_{,\{z_j\}}$ and let

$$\boldsymbol{I} = \begin{pmatrix} O & \boldsymbol{E} \\ -\boldsymbol{E} & O \end{pmatrix}, \tag{5.3}$$

where $\boldsymbol{E}$ is the $N \times N$ identity matrix, so that

$$\boldsymbol{I} = -\boldsymbol{I}' = \boldsymbol{I}^{-1}, \qquad \det|\boldsymbol{I}| = 1. \tag{5.4}$$

We may state the fundamental Theorem 2.2 in terms of partition maps as follows:

Theorem 5.2. *A map T of the collection $\mathscr{T}$ is a canonical partition map of multiplier μ if and only if it is a partition map which is such that (Γ_{ij}) satisfies the equation*

$$\sum_j (\Gamma_{ji})' \boldsymbol{I} (\Gamma_{jk}) = \frac{1}{\mu} \begin{pmatrix} \boldsymbol{O} & \delta_{jk} \boldsymbol{E} \\ -\delta_{ji} \boldsymbol{E} & \boldsymbol{O} \end{pmatrix}. \tag{5.5}$$

Proof. Equation (2.48), which defines canonical maps, may be written in the form

$$\vec{L}(\{\tilde{R}\}, \{\tilde{R}\}) = \mu \vec{L}(\{\tilde{R}\}, \{R\}),$$

as follows from part (h) of Theorem 4.3 since we may evaluate the Lagrange bracket vectrix in either the $\{R\}$ coordinates or the $\{\tilde{R}\}$ coordinates. Now, from equations (4.5), (5.3) and the definition of the Lagrange bracket vectrix,

we have

$$\vec{L}(u,v) = \{R\}'_{,u}\,\vec{I}\{R\}_{,v} = \sum_i \{z_i\}'_{,u}\,I\{z_i\}_{,v}\,\vec{e}_i$$

so that

$$\begin{aligned}\vec{L}(\{Z_i\},\{Z_k\}) &= (\{z_j\}_{,\{Z_i\}})'\,I\,(\{z_j\}_{,\{Z_k\}})\\ &= (\Gamma_{ji})'\,I\,(\Gamma_{jk})\,\vec{e}_j.\end{aligned}$$

Thus, evaluating in the $\{Z_i\}$ coordinates, we have

$$\begin{aligned}(\Gamma_{ji})'\,I\,(\Gamma_{jk})\,\vec{e}_j &= \frac{1}{\mu}\begin{pmatrix} E & O \\ O & \delta_{ji}E\end{pmatrix} I \begin{pmatrix} E & O \\ O & \delta_{jk}E\end{pmatrix}\vec{e}_j\\ &= \frac{1}{\mu}\begin{pmatrix} O & \delta_{jk}E \\ -\delta_{jk}E & O\end{pmatrix}\vec{e}_j\end{aligned}$$

upon requiring the map to be a partition map. Q.E.D.

Generating Vectors for Canonical Partition Maps

We now establish a collection of canonical partition maps which will be of general use. From the proof of Theorem 3.4, it was shown that a map was canonical if and only if

$$\mu(p_{\alpha i}\,\mathrm{q}_{\alpha,i} - H) - P_{\beta j}\,Q_{\beta,j} + \mathrm{K}$$

differed from zero by an element of the null set of the Euler-Lagrange operator,

$$\mathcal{N}(\varepsilon) = \sum_{s=0}^{r} O_s(B_{\alpha i}\,a_{\alpha,i}),$$

characterized by the equations

$$B_{,a_\alpha} = \partial_i B_{\alpha i}, \qquad B_{\alpha i, a_\beta} = B_{\beta i, a_\alpha}$$

and all other B's identically zero. Since the above equations imply that there exists an n-vector Γ such that

$$B = \partial_i \Gamma_i, \qquad B_{\alpha i} = \Gamma_{i, a_\alpha},$$

it follows that

$$\mathcal{N}(\varepsilon) = \partial_i \Gamma_i + \Gamma_{i, a_\alpha}\,a_{\alpha,i} = \Gamma_{i,i}.$$

If $\det|\boldsymbol{A}|$, where $\boldsymbol{A}$ is defined by equation (2.16), is non-zero, we may solve $\mathrm{q}_\alpha(P, Q, \boldsymbol{x})$ for Q_β so as to obtain $Q_\beta = Q_\beta(P, \mathrm{q}, \boldsymbol{x})$. Set

$$\Gamma_{i,i}(P, Q, \boldsymbol{x}) + (P_{\beta j}\,Q_\beta)_{,j} = W_{i,i}(P, \mathrm{q}, \boldsymbol{x})$$

where Q_β is written in terms of $(P, \mathrm{q}, \boldsymbol{x})$ by $Q_\beta = Q_\beta(P, \mathrm{q}, \boldsymbol{x})$. Thus

$$\begin{aligned}\Gamma_{i,i} &= W_{i,i}(P, \mathrm{q}, \boldsymbol{x}) - (P_{\beta j}\,Q_\beta)_{,j}\\ &= W_{i,\mathrm{q}_\alpha}\,\mathrm{q}_{\alpha,i} + W_{i,P_{\beta j}}\,P_{\beta j,i} + \partial_i W_i - Q_\beta\,P_{\beta j,j} - P_{\beta j}\,Q_{\beta,j},\end{aligned}$$

and hence

$$\begin{aligned}&\mu(p_{\alpha i}\,\mathrm{q}_{\alpha,i} - H) - P_{\beta j}\,Q_{\beta,j} + \mathrm{K}\\ &\quad = W_{i,\mathrm{q}_\alpha}\,\mathrm{q}_{\alpha,i} + W_{i,P_{\beta j}}\,P_{\beta j,i} + \partial_i W_i - Q_\beta\,P_{\beta j,j} - P_{\beta j}\,Q_{\beta,j}.\end{aligned}$$

If this is to be satisfied for arbitrary (P, q) independent of H, it is required that

$$\mathrm{K} = \partial_i W_i + \mu H,$$

$$\mu p_{\alpha i} = W_{i, \mathrm{q}_\alpha},$$

$$Q_\beta \delta_{ij} = W_{i, P_{\beta j}}.$$

The last of these equations implies

$$W_{i, P_{\beta j}} = 0, \qquad i \neq j$$

$$\sum_i W_{i, P_{\beta i}} = \sum_j W_{j, P_{\beta j}}.$$

We have thus proved the following

Theorem 5.3. *Let (p, q) form a Hamiltonian system base H and T be a map of the collection $\mathcal{T}$; then T is a canonical map if there exists a vectrix function $\vec{W}(P, \mathrm{q}, \boldsymbol{x})$ of class C^2 such that*

$$\mathrm{K} = \partial_i W_i + \mu H, \tag{5.6}$$

$$\mu p_{\alpha i} = W_{i, \mathrm{q}_\alpha}, \tag{5.7}$$

$$Q_\beta = \frac{1}{n} W_{i, P_{\beta i}}, \tag{5.8}$$

$$W_{i, P_{\beta j}} = 0, \qquad i \neq j, \tag{5.9}$$

$$\sum_i W_{i, P_{\beta i}} = \sum_j W_{j, P_{\beta j}}. \tag{5.10}$$

That equations (5.6) through (5.10) determine the map T is seen from the fact that since $\det |\boldsymbol{A}|$ was assumed nonzero, we may solve equation (5.8) for $\mathrm{q}_\alpha(P, Q, \boldsymbol{x})$, which when substituted into equation (5.7) gives $p_{\alpha i}(P, Q, \boldsymbol{x})$.

Extended Point Transformations

Theorem 5.3 forms the basis for the development of a collection of canonical partition maps which are the canonical extension of generalized point transformations.

Definition. **A generalized point transformation** is a transformation which is representable by the equations of transformation $Q = Q(\mathrm{q}, \boldsymbol{x})$ where $\det (Q_{\beta, \mathrm{q}_\alpha})$ is nonzero for all $\boldsymbol{x}$ in a given $\mathcal{D}_n^*$.

Theorem 5.4. *To every generalized point transformation there corresponds a nondenumerable collection of extensions such that the resulting transformation equations are canonical. In particular, if $Q = Q(\mathrm{q}, \boldsymbol{x})$ is a given generalized point transformation, then*

$$Q_\beta = Q_\beta(\mathrm{q}, \boldsymbol{x}), \tag{5.11}$$

$$P_{\beta i} = (Q_{\beta, \mathrm{q}_\alpha})^{-1} (\mu p_{\alpha i} + S_{i, \mathrm{q}_\alpha}) \tag{5.12}$$

is canonical, where S_i are the components of an arbitrary n-vector function of $\mathrm{q}, \boldsymbol{x}$ which is of class C^2 in q and of class C^1 in $\boldsymbol{x}$.

Proof. From equations (5.8) through (5.10) of Theorem 5.3, we have

$$Q_\beta(\mathrm{q}, \boldsymbol{x}) = \frac{1}{n} W_{i, P_{\beta i}}, \quad (P, \mathrm{q}, \boldsymbol{x}) = \sum_j W_{j, P_{\beta j}}, \quad W_{i, P_{\beta j}} = 0, \qquad i \neq j.$$

Let $Q_\beta = Q_\beta(\mathrm{q}, \boldsymbol{x})$ be the given generalized point transformation; then identifying this Q with the Q in Theorem 5.3 and solving for W_i results in

$$W_i = Q_\beta(\mathrm{q}, \boldsymbol{x})\, P_{\beta i} - S_i(\mathrm{q}, \boldsymbol{x})$$

where S_i is an arbitrary vector function of class C^2 in q and of class C^1 in $\boldsymbol{x}$. Substituting this expression into equation (5.7) gives

$$\mu\, p_{\alpha i} = W_{i, \mathrm{q}_\alpha} = Q_{\beta, \mathrm{q}_\alpha} P_{\alpha i} - S_{i, \mathrm{q}_\alpha}, \tag{5.13}$$

and hence

$$P_{\beta i} = (Q_{\beta, \mathrm{q}_\alpha})^{-1} (\mu\, p_{\alpha i} + S_{i, \mathrm{q}_\alpha}),$$

which is well defined owing to the fact that $\det\left[(Q_{\beta, \mathrm{q}_\alpha})\right] \neq 0$ since the original transformation was assumed to be a generalized point transformation. Q. E. D.

Henceforth, we shall refer to equations (5.11) and (5.12) as the canonical extensions of generalized point transformations.

The properties of the canonical extension of generalized point transformations, given by the last theorem, are summarized in the following theorems.

Theorem 5.5. *The canonical extensions of generalized point transformations are canonical partition maps.*

Proof. This result follows immediately from the definition of canonical partition maps and equation (5.13).

Theorem 5.6. *The collection of all canonical extensions of all generalized point transformations forms a group.*

Proof. From equation (5.11) it is seen that the matrix $\boldsymbol{B}$ defined by equation (2.17) is a zero matrix. Hence, the product of the associated Jacobian matrices of the canonical extensions of two generalized point transformations is a Jacobian matrix of a canonical extension of a generalized point transformation. Since this was the only result which was lacking in order that the collection of all canonical partition maps form a group, the results follows.

Corollary. *The collection of all canonical partition maps possesses a sub-collection which forms a group, namely the canonical extensions of generalized point transformations.*

The canonical map used by LANCZOS[18] to examine the electric field effects in his unified field theory is in actuality a canonical partition map as is evident by inspection of equation (6.8) of his cited paper. This is one of the most vivid examples of the use of generalized canonical maps to yield fundamental information about the structure and implications of a system of partial differential equations of the generalized Hamiltonian form. It is one of the few, if not the only example of generalized canonical maps in the present literature.

---

[18] LANCZOS, C.: Rev. Mod. Phys. **29**, No. 3, 341—342 (July 1957).

Section VI. Infinitesimal Canonical Maps

Infinitesimal Maps

Consider those elements of the collection $\mathcal{T}$ which are such that a constant multiple of the image of $\{R\}$ differs from $\{R\}$ by a quantity $\{\Delta R\}$ which is an $N(n+1)$-element column matrix whose elements are of infinitesimal order ε; that is,

$$\lambda\{\tilde{R}\} = \{R\} + \{\Delta R\}, \tag{6.1}$$

where λ is a constant.

Definition. A map T of the collection $\mathcal{T}$ will be said to be an **infinitesimal map with dilatation** λ if the associated transformation equations satisfy the equation (6.1) for $\{\Delta R\}$ of infinitesimal order.

Definition. A map of the collection $\mathcal{T}$ will be said to be an **infinitesimal map without dilatation,** or more simply an infinitesimal map, if $\lambda = 1$. Computing the matrix M^{-1} from equation (6.1), we have

$$\lambda\{\tilde{R}\}_{,\{R\}} = \boldsymbol{E}^* + \{\Delta R\}_{,\{R\}}$$

so that

$$\boldsymbol{M}^{-1} = \frac{1}{\lambda}(\boldsymbol{E}^* + \varepsilon\,\boldsymbol{m}), \qquad \varepsilon\,\boldsymbol{m} = \{\Delta R\}_{,\{R\}}, \tag{6.2}$$

which implies

$$\boldsymbol{M} = \lambda(\boldsymbol{E}^* - \varepsilon\,\boldsymbol{m}) + \star \tag{6.3}$$

to within a matrix whose elements are of order ε^2, denoted by $\star$.

Let us now restrict our attention to those infinitesimal maps with dilatation λ which are canonical (*i.e.*, $\boldsymbol{M}$ is symplectic with multiplier μ). Using the results of Theorem 4.2, we may write $\boldsymbol{M} = \tilde{\boldsymbol{M}}\boldsymbol{M}^*$, which from the definition of $\tilde{\boldsymbol{M}}$, namely $\tilde{\boldsymbol{M}} = \frac{1}{\sqrt{\mu}}\boldsymbol{E}^*$, gives $\boldsymbol{M} = \frac{1}{\sqrt{\mu}}\boldsymbol{M}^*$. Comparing this with equation (6.3) yields

$$\boldsymbol{M}^* = \boldsymbol{E}^* - \varepsilon\,\boldsymbol{m},$$

$$\mu = 1/\lambda^2,$$

from which we may conclude that a canonical map which is infinitesimal with dilatation λ is such that its multiplier is the reciprocal square of the dilatation. We also know from Theorem 4.2 that $\boldsymbol{M}^*$ is a basic symplectic matrix, and hence $\boldsymbol{M}^*$ is the associated Jacobian matrix of an infinitesimal map since from the above considerations $\mu = 1$ which implied that $\lambda = 1$.

We consider, from this point on, only infinitesimal maps since infinitesimal maps with dilatation can be constructed from infinitesimal maps by an appropriate $\tilde{\boldsymbol{M}}$.

If an infinitesimal map is to be canonical, it is both necessary and sufficient, by Theorem 2.2, to require the associated Jacobian matrix of the map to satisfy

$$\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} = \vec{\boldsymbol{I}},$$

since infinitesimal canonical maps are elements of the basic symplectic group. Substituting from equation (6.3), where λ is unity, gives

$$\boldsymbol{M}'\vec{\boldsymbol{I}}\boldsymbol{M} = \vec{\boldsymbol{I}} - \varepsilon(\vec{\boldsymbol{I}}\boldsymbol{m} + \boldsymbol{m}'\vec{\boldsymbol{I}}) + \star,$$

and hence

$$-\varepsilon(\vec{\boldsymbol{I}}\boldsymbol{m} + \boldsymbol{m}'\vec{\boldsymbol{I}}) + * = \vec{\boldsymbol{O}}$$

so that, to the first order in ε, it is both necessary and sufficient for the matrix $\boldsymbol{m}$ to satisfy

$$\vec{\boldsymbol{I}}\boldsymbol{m} + \boldsymbol{m}'\vec{\boldsymbol{I}} = \vec{\boldsymbol{O}}$$

in order for the infinitesimal map to be canonical. Noting that $\boldsymbol{m}'\vec{\boldsymbol{I}} = -(\vec{\boldsymbol{I}}\boldsymbol{m})'$, the following theorem results:

***Theorem 6.1.** *An infinitesimal map of the collection $\mathscr{T}$ is an infinitesimal canonical map (of the basic symplectic group) if and only if the vectrix $\vec{\boldsymbol{I}}\boldsymbol{m}$ is symmetric.*

Analytic Characterization and Generating Vectrices

Theorem 6.1 provides the basis for the explicit analytic characterization of the collection of infinitesimal canonical maps of the unitary symplectic group. Differentiating equation (6.1) with respect to $\{R\}$ gives

$$\{\widetilde{R}\}_{,\{R\}} = \boldsymbol{E}^* + \{\Delta R\}_{,\{R\}} = \boldsymbol{M}^{-1},$$

which when compared with equation (6.2) results in

$$\varepsilon\,\boldsymbol{m} = \{\Delta R\}_{,\{R\}}.$$

Let

$$\{\Delta R\} \overset{\text{def}}{=} \varepsilon\{\delta R\}, \tag{6.4}$$

and hence

$$\boldsymbol{m} = \{\delta R\}_{,\{R\}}. \tag{6.5}$$

Now

$$\vec{\boldsymbol{I}}\boldsymbol{m} = \begin{pmatrix} (\delta p_{\alpha i,\mathrm{q}_\beta})\vec{e}_i & (\delta p_{\alpha i,p_{\beta 1}})\vec{e}_i & \cdots & (\delta p_{\alpha i,p_{\beta n}})\vec{e}_i \\ (-\delta \mathrm{q}_{\alpha,\mathrm{q}_\beta})\vec{e}_1 & -(\delta \mathrm{q}_{\alpha,p_{\beta 1}})\vec{e}_1 & \cdots & -(\delta \mathrm{q}_{\alpha,p_{\beta n}})\vec{e}_1 \\ \vdots & \vdots & \ddots & \vdots \\ -(\delta \mathrm{q}_{\alpha,\mathrm{q}_\beta})\vec{e}_n & -(\delta \mathrm{q}_{\alpha,p_{\beta 1}})\vec{e}_n & \cdots & -\delta(\mathrm{q}_{\alpha,p_{\beta n}})\vec{e}_n \end{pmatrix},$$

as is seen by direct expansion of $\vec{\boldsymbol{I}}\boldsymbol{m}$ using equation (6.5), and

$$\boldsymbol{m}'\vec{\boldsymbol{I}}' = \begin{pmatrix} (\delta p_{\alpha i,\mathrm{q}_\beta})'\vec{e}_i & -(\delta \mathrm{q}_{\alpha,\mathrm{q}_\beta})'\vec{e}_1 & \cdots & -(\delta \mathrm{q}_{\alpha,\mathrm{q}_\beta})'\vec{e}_n \\ (\delta p_{\alpha i,p_{\beta 1}})'\vec{e}_i & -(\delta \mathrm{q}_{\alpha,\mathrm{q}_{\beta 1}})'\vec{e}_1 & \cdots & -(\delta \mathrm{q}_{\alpha,p_{\beta 1}})\vec{e}_n \\ \vdots & \vdots & & \\ (\delta p_{\alpha i,p_{\beta n}})'\vec{e}_i & -(\delta \mathrm{q}_{\alpha,\mathrm{q}_{\beta n}})'\vec{e}_1 & \cdots & -(\delta \mathrm{q}_{\alpha,p_{\beta n}})'\vec{e}_n \end{pmatrix}.$$

If $\vec{\boldsymbol{I}}\boldsymbol{m}$ $_{-}\boldsymbol{m}'\vec{\boldsymbol{I}}'$, so as to satisfy Theorem 6.1, it is required that each component of this vectrix equation be satisfied (that is, the matrices which multiply $\vec{e}_k$, as a result of the equation $\vec{\boldsymbol{I}}\boldsymbol{m} - \boldsymbol{m}'\vec{\boldsymbol{I}}' = \boldsymbol{O}$, must be equal). Thus the matrix

which multiplies $\vec{e}_k$ gives for $k=1, \ldots, n$

$$\delta p_{\alpha k, q_\beta} = \delta p_{\beta k, q_\alpha}, \tag{6.6}$$

$$\delta q_{\alpha, p_{\beta k}} = \delta q_{\beta, p_{\alpha k}}, \tag{6.7}$$

$$-\delta q_{\alpha, q_\beta} = \sum_k \delta p_{\beta k, p_{\alpha k}}, \tag{6.8}$$

$$\left.\begin{aligned} \delta p_{\alpha k, p_{\beta l}} &= 0 \quad (6.9) \\ \delta q_{\alpha, p_{\beta l}} &= 0 \quad (6.10) \end{aligned}\right\} \quad l \neq k.$$

Since the above equations must be satisfied for all $k=1, \ldots, n$, equation (6.10) implies that δq_α does not depend on $p_{\beta j}$ for all β and j. Hence, equation (6.7) is identically satisfied by the conditions imposed by equation (6.10). Thus, since δq_α can depend only on q_β and $\boldsymbol{x}$ at most, equation (6.8) implies that $\delta p_{\beta k}$ must be linear in the $p_{\alpha i}$. However, equation (6.9) implies that $\delta p_{\alpha i}$ can depend at most on $p_{\alpha i}$, q_α, and $\boldsymbol{x}$ for fixed i. Hence, equations (6.6) through (6.10) reduce to

$$\delta q_\alpha = \delta q_\alpha(q_\beta, \boldsymbol{x}), \tag{6.11}$$

$$\delta p_{\alpha i} = h_{\alpha\gamma}(q_\beta, \boldsymbol{x})\, p_{\gamma i} + g_{\alpha i}(q_\beta, \boldsymbol{x}), \tag{6.12}$$

$$\delta p_{\alpha i, q_\beta} = \delta p_{\beta i, q_\alpha}, \tag{6.13}$$

$$-\delta q_{\alpha, q_\beta} = \sum_k \delta p_{\beta k, q_{\alpha k}}, \tag{6.14}$$

where $h_{\alpha\gamma}$ and $g_{\alpha i}$ are arbitrary collections of functions of class C^1. Substituting equation (6.12) into equation (6.14) gives

$$-\delta q_{\alpha, q_\beta} = h_{\beta\alpha}(q_\gamma, \boldsymbol{x}), \tag{6.15}$$

which are completely integrable if and only if

$$h_{\beta\alpha, q_\gamma} = h_{\gamma\alpha, q_\beta}. \tag{6.16}$$

Substituting equation (6.12) into equation (6.13) results in

$$h_{\alpha\gamma, q_\beta} p_{\alpha i} + g_{\alpha i, q_\beta} = h_{\beta\gamma, q_\alpha} p_{\gamma i} + g_{\beta i, q_\alpha}$$

which, upon using the integrability conditions for equation (6.15) (that is, equation (6.16)), will be satisfied if

$$g_{\alpha i, q_\beta} = g_{\beta i, q_\alpha}. \tag{6.17}$$

Equations (6.16) and (6.17), which are necessary conditions for general solutions of equations (6.13) and (6.14) to exist, are satisfied by

$$h_{\alpha\gamma} = f_{\gamma, q_\alpha}(q, \boldsymbol{x}), \tag{6.18}$$

$$g_{\alpha i} = I_{i, q_\alpha}(q, \boldsymbol{x}), \tag{6.19}$$

where f_γ and I_{i, q_α} are arbitrary functions of class C^2 in q and of class C^1 in $\boldsymbol{x}$. Substituting equation (6.18) into equation (6.15) gives

$$-\delta q_{\alpha, q_\beta} = f_{\alpha, q_\beta},$$

which upon integration results in

$$\delta q_\alpha = -f_\alpha(q, \boldsymbol{x}).$$

Theorem 6.2. *For the cases $n>1$ an infinitesimal map of the collection $\mathscr{T}$ is an infinitesimal canonical map if and only if*

$$\Delta q_\alpha = -\varepsilon f_\alpha(q, \boldsymbol{x}), \tag{6.20}$$

$$\Delta p_{\alpha i} = f_{\gamma, q_\alpha}(q, \boldsymbol{x})\, p_{\gamma i} + I_{i, q_\alpha}(q, \boldsymbol{x}) \tag{6.21}$$

where $f_\alpha(q, \boldsymbol{x})$ and $I_i(q, \boldsymbol{x})$ are arbitrary functions of class C^2 in q *and of class C^1 in $\boldsymbol{x}$.*

Proof. For the case $n>1$ equations (6.20) and (6.21) follow directly from equations (6.4), (6.12), (6.18), and (6.19), which constitute a general solution to the symmetry condition of the vectrix $\vec{\boldsymbol{Im}}$ as given in Theorem 6.1. That the above equations do not apply in the case $n=1$ is seen from the fact that equations (6.9) and (6.10) are no longer required since k can have only the value unity, and equations (6.6), (6.7), and (6.8) become

$$\delta p_{\alpha 1, q_\beta} = \delta p_{\beta 1, q_\alpha},$$

$$\delta q_{\alpha, p_{\beta 1}} = \delta q_{\beta, p_{\alpha 1}},$$

$$-\delta q_{\alpha, q_\beta} = \delta p_{\beta 1, p_{\alpha 1}}$$

which are satisfied by

$$\delta p_{\alpha i} = -W_{, q_\alpha},$$

$$\delta q_\alpha = W_{, p_{\alpha 1}},$$

where W is an arbitrary function of $(p_{\alpha i}, q_\alpha, \boldsymbol{x})$ of class C^2; the present results thus reduce to those of the classic literature for the case $n=1$. Q.E.D.

Let

$$\vec{W} = \vec{e}_i\left(f_\gamma(q, \boldsymbol{x})\, p_{\gamma i} + I_i(q, \boldsymbol{x})\right); \tag{6.22}$$

then equations (6.20) and (6.21) are equivalent to

$$\{\Delta R\} = \varepsilon\, \vec{\boldsymbol{I}}_l^{-1} \bullet \{\vec{W}_{,R}\}, \tag{6.23}$$

as is seen by direct expansion. Thus equation (6.1) becomes

$$\{\widetilde{R}\} = \{R\} + \varepsilon\, \vec{\boldsymbol{I}}_l^{-1} \bullet \{\vec{W}_{,R}\}. \tag{6.24}$$

The vectrix function $\vec{W}$ will be referred to as the **generating vectrix** of the infinitesimal canonical map given by equation (6.1).

It now remains to determine the remainder function associated with the infinitesimal canonical map specified by equation (6.24). By equation (2.22) the remainder function U satisfied the equation

$$\vec{\boldsymbol{I}} \bullet \underset{0}{\vec{\nabla}}\{\widetilde{R}\} = \{U_{,\widetilde{R}}\}.$$

From equation (6.1)

$$\underset{0}{\vec{\nabla}}\{\widetilde{R}\} = \underset{0}{\vec{\nabla}}\{R\} + \underset{0}{\vec{\nabla}}\{\Delta R\} = \underset{0}{\vec{\nabla}}\{\Delta R\} = \varepsilon\, \underset{0}{\vec{\nabla}}\left(\boldsymbol{I}_l^{-1} \bullet \{W_{,R}\}\right),$$

and since

$$\{U_{,\tilde{R}}\} = \boldsymbol{M}' \{U_{,R}\} = (\boldsymbol{E}^* - \varepsilon \boldsymbol{m}') \{U_{,R}\},$$

it is thus required that

$$\{U_{,R}\} = \varepsilon(\vec{\boldsymbol{I}} \bullet \underset{0}{\vec{\nabla}}) (\vec{\boldsymbol{I}}_l^{-1} \bullet \{\vec{W}_{,R}\}),$$

to within a quantity of order ε^2 $(i.e., O(U) = \varepsilon)$,

$$\{U_{,R}\} = \varepsilon \underset{0}{\vec{\nabla}} \bullet \vec{\boldsymbol{I}} (\vec{\boldsymbol{I}}_l^{-1} \bullet \{\vec{W}_{,R}\})$$

which shows, from the form of $\vec{W}$, that we may take $\{\vec{C}\} = 0$ in Theorem 1.2 with no loss of generality so as to obtain

$$\{U_{,R}\} = \varepsilon \underset{0}{\vec{\nabla}} \bullet \{\vec{W}_{,R}\} = \varepsilon \{(\underset{0}{\vec{\nabla}} \bullet \vec{W})_{,R}\},$$

and hence

$$U = \varepsilon \underset{0}{\vec{\nabla}} \bullet \vec{W}. \tag{6.25}$$

Thus, by equation (2.47), we have

$$\mathrm{K} = H - \varepsilon \underset{0}{\vec{\nabla}} \bullet \vec{W}. \tag{6.26}$$

From inspection of equations (6.20) and (6.21) we have

Theorem 6.3. *The most general infinitesimal canonical map, for the case $n > 1$, is a canonical partition map which is the extension of an infinitesimal generalized point transformation.*

Hamilton-Jacobi Theory for the Case $n = 1$

The true nature of the generalization over the classic theory $(n = 1)$ is pointed up most graphically by contrasting a particular result obtained in the case $n = 1$ with the lack of any counterpart when $n > 1$. For $n = 1$, from equations (6.6) and (6.10) we have

$$\begin{aligned} \Delta p_{\alpha 1} &= -\varepsilon W_{,\mathrm{q}_\alpha}, \\ \Delta \mathrm{q}_\alpha &= \varepsilon W_{,p_{\alpha 1}} \end{aligned} \tag{6.27}$$

and the corresponding Hamiltonian equations

$$\begin{aligned} \dot{p}_{\alpha 1} &= -H_{,\mathrm{q}_\alpha}, \\ \dot{\mathrm{q}}_\alpha &= H_{,p_{\alpha 1}}, \end{aligned} \qquad (\dot{\ }) = \frac{d}{d x_1}. \tag{6.28}$$

If H is of class C^2, we may identify εW with $H \Delta x_1$, since in the case $n = 1$ there is no restriction on the form of W other than that it be of class C^2 in (p, q). Thus, equations (6.27) become

$$\begin{aligned} \Delta p_{\alpha 1} &= -H_{,\mathrm{q}_\alpha} \Delta x_1 \\ \Delta \mathrm{q}_\alpha &= H_{,p_{\alpha 1}} \Delta x_1, \end{aligned} \tag{6.29}$$

which, by equations (6.28), become

$$\begin{aligned} \Delta p_{\alpha 1} &= \dot{p}_{\alpha 1} \Delta x_1, \\ \Delta \mathrm{q}_\alpha &= \dot{\mathrm{q}}_\alpha \Delta x_1, \end{aligned} \tag{6.30}$$

and hence

$$\begin{aligned} \tilde{p}_{\alpha 1} &= p_{\alpha 1} + \dot{p}_{\alpha 1} \Delta x_1, \\ \tilde{q}_\alpha &= q_\alpha + \dot{q}_\alpha \Delta x_1. \end{aligned} \tag{6.31}$$

If $(p_{\alpha 1}, q_\alpha)$ are identified with initial conditions, equations (6.31) state that the solution to equations (6.28), for sufficiently small Δx_1, is given by an infinitesimal canonical map with generating function $H \Delta x_1$, since equations (6.31) are just the first two terms in the power series expansion of the solution in a neighborhood of the initial conditions. Thus, in the case $n = 1$, the solution to equations (6.28) is the result of successive applications of infinitesimal canonical maps. This fact coupled with the group property of canonical maps states that the solution of any problem described by equations (6.28) is the unfolding of a canonical map from the initial conditions to the final state, at some time $\underset{1}{x_1}$ say. In other words, there exists a canonical map which maps the initial conditions into the solution for a fixed value of the independent variable x_1 in the above case $\underset{1}{x_1}$. It is this result which forms the basis for the integration of equations (6.28) in the case $n = 1$.

On the Existence of a Generalization of Hamilton-Jacobi Theory for Case $n > 1$

Let us examine the case where $n > 1$ to see whether there exist bases functions such that the same result follows under an infinitesimal map as shown above for the case $n = 1$. From equations (6.22) and (6.23) we have

$$\Delta q_\alpha \doteq -\varepsilon \sum_i W_{i, p_{\alpha i}} = -\varepsilon \sum_j W_{j, p_{\alpha j}}$$

for all i and j equal 1 to n. In order that Δq_α will be equal to $d q_\alpha$, it is required that

$$-\varepsilon \sum_i W_{i, p_{\alpha i}} = q_{\alpha, i} \, d x^i, \qquad d x^i \, d x^i = \varepsilon^2. \tag{6.32}$$

From equation (2.1), since (p, q) are assumed to form a Hamiltonian system with base H, we have

$$q_{\alpha, i} = H_{, p_{\alpha i}},$$

so that we require

$$-\varepsilon \sum_i W_{i, p_{\alpha i}} = H_{, p_{\alpha k}} \, d x^k. \tag{6.33}$$

Since $W_{i, p_{\alpha i}}$ must satisfy the relation

$$\sum_i W_{i, p_{\alpha i}} = \sum_j W_{j, p_{\alpha j}}, \tag{6.34}$$

we obtain, upon integrating equation (6.33) subject to the conditions (6.34), the most general form which H can have in order for there to exist a solution to equation (6.33) under the conditions (6.34), namely

$$H = -\psi_{\alpha i}(q, \boldsymbol{x}) \, p_{\alpha i} + \eta(q, \boldsymbol{x}). \tag{6.35}$$

In this case

$$\varepsilon W_i = \psi_{\alpha k} \, p_{\alpha i} \, d x_k + \xi_{i j}(q, \boldsymbol{x}) \, d x^j. \tag{6.36}$$

From equation (6.7) we have $\Delta p_{\beta j} = \varepsilon W_{j, q_\beta}$,

which, by the use of equation (6.36) becomes

$$\Delta p_{\beta j} = (\psi_{\alpha k, q_\beta} p_{\alpha j} + \xi_{jk, q_\beta})\, d x^k. \tag{6.37}$$

We have

$$d p_{\beta j} = p_{\beta j, k}\, d x^k,$$

so that in order that $d p_{\beta j}$ be equal to $\Delta p_{\beta j}$ we require

$$p_{\beta j, k}\, d x^k = (\psi_{\alpha k, q_\beta} p_{\alpha j} + \xi_{jk, q_\beta})\, d x^k$$

which for arbitrary $d x^k$ gives

$$p_{\beta j, k} = \psi_{\alpha k, q_\beta} p_{\alpha j} + \xi_{jk, q_\beta}. \tag{6.38}$$

If we set j equal to k and sum over the common index and require that the arbitrary functions ξ_{jk} satisfy

$$\xi_{ii} = -\eta, \tag{6.39}$$

we obtain, using equation (6.34),

$$p_{\beta i, i} = -H_{, q_\beta}$$

which is just equation $(2.1)_1$. Thus, equation (6.38) is consistent with the Hamiltonian equations which (p, q) satisfy if equation (6.39) is satisfied. Hence, we have

$$d q_\alpha = \Delta q_\alpha = -\psi_{\alpha k}\, d x^k, \tag{6.40}$$

$$d p_{\beta j} = \Delta p_{\beta j} = (\psi_{\alpha k} p_{\alpha j} + \xi_{jk})_{, q_\beta}\, d x^k, \tag{6.41}$$

$$\xi_{ii} = -\eta, \tag{6.42}$$

$$H = -\psi_{\alpha i}(q, \boldsymbol{x})\, p_{\alpha i} + \eta(q, \boldsymbol{x}). \tag{6.43}$$

Thus far we have satisfied the condition $d q_\alpha = \Delta q_\alpha$, $d p_{\beta j} = \Delta p_{\beta j}$ in a formal manner only. It remains to show that the assumed differentials are actually differentials. From equation (6.40) we have

$$d q_\alpha = q_{\alpha, k}\, d x^k = -\psi_{\alpha k}\, d x^k, \tag{6.44}$$

which yields $q_{\alpha, k} = -\psi_{\alpha k}$. If $d q_\alpha$ is to be a differential (exact), it is required that

$$\psi_{\alpha k, j} = \psi_{\alpha j, k}. \tag{6.45}$$

Since

$$\psi_{\alpha k, j} = \psi_{\alpha k, q_\beta} q_{\beta, j} + \partial_j \psi_{\alpha k} = -\psi_{\alpha j, q_\beta} \psi_{\beta k} + \partial_k \psi_{\alpha j},$$

by equation (6.45), we require $\psi_{\alpha k}$ to be such that

$$-\psi_{\alpha k, q_\beta} \psi_{\beta j} + \partial_i \psi_{\alpha k} = -\psi_{\alpha j, q_\beta} \psi_{\beta k} + \partial_k \psi_{\alpha j}. \tag{6.46}$$

Similarly, from equation (6.41) we have

$$d p_{\beta j} = p_{\beta j, k}\, d x^k = (\psi_{\alpha k} p_{\alpha j} + \xi_{jk})_{, q_\beta}\, d x^k \tag{6.47}$$

which yields

$$p_{\beta j, k} = (\psi_{\alpha k} p_{\alpha j} + \xi_{jk})_{, q_\beta}. \tag{6.48}$$

Hence, if $d p_{\beta j}$ is to be an exact differential, it is required that

$$(\psi_{\alpha k,\mathrm{q}_\beta} p_{\alpha j} + \xi_{jk,\mathrm{q}_\beta})_{,l} = (\psi_{\alpha l,\mathrm{q}_\beta} p_{\alpha j} + \xi_{jl,\mathrm{q}_\beta})_{,k}.$$

But

$$\begin{aligned}(\psi_{\alpha k,\mathrm{q}_\beta} p_{\alpha j} + \xi_{jk,\mathrm{q}_\beta})_{,l} = {} & - \psi_{\alpha k,\mathrm{q}_\beta \mathrm{q}_\gamma} p_{\alpha j} \psi_{\gamma l} - \\ & - \xi_{jk,\mathrm{q}_\beta \mathrm{q}_\gamma} \psi_{\gamma l} + \\ & + \psi_{\alpha k,\mathrm{q}_\beta} (\psi_{\gamma l,\mathrm{q}_\alpha} p_{\gamma j} + \xi_{jl,\mathrm{q}_\alpha}) + \\ & + \partial_l (\psi_{\alpha k,\mathrm{q}_\beta} p_{\alpha j} + \xi_{jk,\mathrm{q}_\beta})\end{aligned}$$

by equations (6.44) and (6.48). Thus, ψ and ξ must be such that equations (6.46) and the following are satisfied:

$$\begin{aligned}0 = {} & p_{\alpha j} [\psi_{\alpha k,\mathrm{q}_\gamma} \psi_{\gamma l} - \psi_{\alpha l,\mathrm{q}_\gamma} \psi_{\gamma k} + \partial_l \psi_{\alpha k} - \partial_k \psi_{\alpha l}]_{,\mathrm{q}_\beta} + \\ & + [\psi_{\alpha k} \xi_{jl,\mathrm{q}_\alpha} - \psi_{\alpha l} \xi_{jk,\mathrm{q}_\alpha} + \partial_l \xi_{jk} - \partial_k \xi_{jl}]_{,\mathrm{q}_\beta}.\end{aligned} \tag{6.49}$$

By equation (6.46) we have

$$\psi_{\alpha k,\mathrm{q}_\gamma} \psi_{\gamma l} - \psi_{\alpha l,\mathrm{q}_\gamma} \psi_{\gamma k} = \partial_l \psi_{\alpha k} - \partial_k \psi_{\alpha l},$$

so that equation (6.49) becomes

$$\begin{aligned}0 = {} & 2 p_{\alpha j} (\partial_l \psi_{\alpha k} - \partial_k \psi_{\alpha l})_{,\mathrm{q}_\beta} + \\ & + [\psi_{\alpha k} \xi_{jl,\mathrm{q}_\alpha} - \psi_{\alpha l} \xi_{jk,\mathrm{q}_\alpha} + \partial_l \xi_{jk} - \partial_k \xi_{jl}]_{,\mathrm{q}_\beta},\end{aligned} \tag{6.50}$$

which is the second required condition in order for the Hamiltonian system given by equation (6.43) to be completely integrable.

Let λ be a parameter defining a curve lying in the boundary of $\mathscr{D}_n^*$ connecting the points $\boldsymbol{x} = \boldsymbol{a}$ and $\boldsymbol{x} = \boldsymbol{b}$; $\boldsymbol{a}, \boldsymbol{b} \subset \mathscr{D}_n^* \ominus \mathscr{D}_n$. We then have

$$\mathrm{q}_\alpha(\boldsymbol{b}) = \mathrm{q}_\alpha(\boldsymbol{a}) - \psi_{\alpha k} \frac{d x^k}{d\lambda} d\lambda,$$

$$p_{\alpha j}(\boldsymbol{b}) = p_{\alpha j}(\boldsymbol{a}) + (\psi_{\beta k} p_{\alpha j} + \xi_{jk})_{,\mathrm{q}_\alpha} \frac{d x^k}{d\lambda} d\lambda,$$

so that

$$d\mathrm{q}_\alpha / d\lambda = - \psi_{\alpha k}|_{\boldsymbol{x}=\boldsymbol{a}} \hbar^k, \tag{6.51}$$

$$d p_{\alpha j} / d\lambda = (\psi_{\beta k} p_{\beta j} + \xi_{jk})_{,\mathrm{q}_\alpha}|_{\boldsymbol{x}=\boldsymbol{a}} \hbar^k, \tag{6.52}$$

$$d x^k / d\lambda = \hbar^k, \tag{6.53}$$

and hence any initial conditions for which there exists a solution to the Hamiltonian equations with bases function given by equation (6.34) must satisfy equations (6.51), (6.52), and (6.53), since points arbitrarily close together on the boundary can have their respective values of $\{R\}$ connected by a canonical map. We have thus proved the following

Theorem 6.4. *There exists an infinitesimal canonical map which generates a local solution to a generalized Hamiltonian system if and only if the basis function of the system is given by*

$$H = - \psi_{\alpha i}(\mathrm{q}, \boldsymbol{x}) p_{\alpha i} + \eta(\mathrm{q}, \boldsymbol{x}), \tag{6.54}$$

where

$$
\begin{aligned}
-\psi_{\alpha k,\mathrm{q}_\beta}\psi_{\beta j}+\partial_j\psi_{\alpha k} &= -\psi_{\alpha j,\mathrm{q}_\beta}\psi_{\beta k}+\partial_k\psi_{\alpha j},\\
0 &= 2p_{\alpha j}\left[\partial_l\psi_{\alpha k}-\partial_k\psi_{\alpha l}\right]_{,\mathrm{q}_\beta}+\\
&\quad+\left[\psi_{\alpha k}\xi_{jl,\mathrm{q}_\alpha}-\psi_{\alpha l}\xi_{jk,\mathrm{q}_\alpha}+\partial_l\xi_{jk}-\partial_k\xi_{jl}\right]_{,\mathrm{q}_\beta},
\end{aligned}
\tag{6.55}
$$

in which case the generating vector is given by

$$\varepsilon W_i = \psi_{\alpha k}p_{\alpha i}\,d x^k + \xi_{ij}\,d x^j, \tag{6.56}$$

$$\varepsilon^2 = d x^i\, d x^i, \qquad \xi_{ii} = -\eta; \tag{6.57}$$

the increments in (p, q) *are given by*

$$d\mathrm{q}_\alpha = \Delta\mathrm{q}_\alpha = -\psi_{\alpha k}\,d x^k, \tag{6.58}$$

$$d p_{\alpha j} = \Delta p_{\alpha j} = (\psi_{\beta k}p_{\alpha j}+\xi_{jk})_{,\mathrm{q}_\beta}\,d x^k. \tag{6.59}$$

In addition, $p_{\beta j}$ *satisfies the additional equation*

$$p_{\beta j,k} = \psi_{\alpha k,\mathrm{q}_\beta}p_{\alpha j} + \xi_{jk,\mathrm{q}_\beta} \tag{6.60}$$

identically, and the boundary conditions must satisfy

$$
\begin{aligned}
d\mathrm{q}_\alpha/d\lambda &= -\psi_{\alpha k}\,\hbar^k,\\
d p_{\alpha j}/d\lambda &= (\psi_{\beta k}p_{\beta j}+\xi_{jk})_{,\mathrm{q}_\alpha}\,\hbar^k,\\
d x^k/d\lambda &= \hbar^k
\end{aligned}
\tag{6.61}
$$

for all λ, *where* λ *is a geodesic arc parameter in the boundary of* $\mathscr{D}_n^*$, *if solutions are to exist in the large.*

If the above conditions are satisfied, then from the group property of canonical maps we may obtain the solution to the Hamiltonian system at all points of $\mathscr{D}_n^*$. It should be noted that the bases function given by equation (6.54) is such that there does not necessarily exist a corresponding Lagrangean system since the second derivative of H with respect to any $p_{\alpha i}$ is zero (*cf.* Theorem 3.3).

Transformation Groups and Divergence Equations

Let $\{R\}$ be such that $\delta\int_{\mathscr{D}_n^*}\mathscr{L}\,d v_n = 0$, for $\mathscr{L}(R; R_{,i}; \boldsymbol{x})$ given, and consider a transformation on $\{R\}$ defined by

$$\{\widetilde{R}\} = \{R\} + \varepsilon\{f(R)\} + \star, \tag{6.62}$$

$$\vec{V}\{\widetilde{R}\} = \vec{V}\{R\} + \varepsilon\,\vec{V}\{f(R)\} + \star, \tag{6.63}$$

where $\varepsilon \geqq 0$ is a parameter independent of $\{R\}$ and $\boldsymbol{x}$. From the form of equations (6.62) we see that $\varepsilon = 0$ gives the identity map, while

$$\{f\} = \partial\{\widetilde{R}\}/\partial\varepsilon\big|_{\varepsilon=0}. \tag{6.64}$$

Starting with the given $\mathscr{L}(R; R_{,i}; \boldsymbol{x})$, we define a new $L(R; R_{,i}; \boldsymbol{x})$ by

$$L(\widetilde{R}; \widetilde{R}_{,i}; \boldsymbol{x}) = \mathscr{L}(R; R_{,i}; \boldsymbol{x}). \tag{6.65}$$

It should be noted that the function L, as defined by equation (6.65), is not the function $\mathscr{L}(\widetilde{R};\widetilde{R}_{,i};\boldsymbol{x})$. In fact, it is not necessarily true that L is close to $\mathscr{L}(\widetilde{R};\widetilde{R}_{,i};\boldsymbol{x})$ even when ε is very small. We thus consider the problem of what conditions must be placed on the transformation equations in order that L shall be close to $\mathscr{L}(\widetilde{R};\widetilde{R}_{,i};\boldsymbol{x})$, that is,

$$L(\widetilde{R};\widetilde{R}_{,i};\boldsymbol{x})-\mathscr{L}(\widetilde{R};\widetilde{R}_{,i};\boldsymbol{x})=O(\varepsilon). \tag{6.66}$$

Expanding $\mathscr{L}(\widetilde{R};\widetilde{R}_{,i};\boldsymbol{x})$, we have

$$\mathscr{L}(\widetilde{R};\vec{\nabla}\widetilde{R};\boldsymbol{x})=\mathscr{L}(R;\vec{\nabla}R;\boldsymbol{x})+\varepsilon\left[\mathscr{L}_{,\{R\}}\{f\}+\mathscr{L}_{,\vec{\nabla}\{R\}}\bullet\vec{\nabla}\{f\}\right]+\star.$$

Now

$$\vec{\nabla}\bullet\left(\mathscr{L}_{,\nabla\{R\}}\{f\}\right)=\left(\vec{\nabla}\bullet\mathscr{L}_{,\vec{\nabla}\{R\}}\right)\{f\}+\mathscr{L}_{,\vec{\nabla}\{R\}}\bullet\vec{\nabla}\{f\},$$

so that

$$\mathscr{L}(\widetilde{R};\vec{\nabla}\widetilde{R};\boldsymbol{x})=\mathscr{L}(R;\vec{\nabla}R;\boldsymbol{x})+\varepsilon\left[\vec{\nabla}\bullet\left(\mathscr{L}_{,\vec{\nabla}\{R\}}\{f\}\right)+\{E|\mathscr{L}\}_{\{R\}}\{f\}\right]+\star \tag{6.67}$$

since $\{E|\mathscr{L}\}_{\{R\}}=\mathscr{L}_{,\{R\}}-\vec{\nabla}\bullet\mathscr{L}_{,\vec{\nabla}\{R\}}$ by definition. Hence, equation (6.66) becomes

$$L(\widetilde{R};\vec{\nabla}\widetilde{R};\boldsymbol{x})=\mathscr{L}(R;\vec{\nabla}R;\boldsymbol{x})+\varepsilon\left[\vec{\nabla}\bullet\left(\mathscr{L}_{,\vec{\nabla}\{R\}}\{f\}\right)+\{E|\mathscr{L}\}_{\{R\}}\{f\}\right]+\star. \tag{6.68}$$

Since $\{R\}$ was assumed such that $\delta\int\mathscr{L}\,dv_n=0$, we have $\{E|\mathscr{L}\}_{\{R\}}=[0]$; hence a necessary condition that equation (6.66) hold is that $\{f\}$ be such that

$$\vec{\nabla}\bullet\left(\mathscr{L}_{,\vec{\nabla}\{R\}}\{f\}\right)=0. \tag{6.69}$$

From equation (6.67), we also have that

$$\mathscr{L}(\widetilde{R};\vec{\nabla}\widetilde{R};\boldsymbol{x})=\mathscr{L}(R;\vec{\nabla}R;\boldsymbol{x})+\star$$

(that is, $\mathscr{L}$ is invariant to within terms of order ε^2) if equation (6.69) is satisfied and conversely. We have thus proved the following well known

Theorem 6.5. *If $\mathscr{L}(R;\vec{\nabla}R;\boldsymbol{x})$ is invariant under the transformations* (6.62), (6.63) *to within terms of order ε^2, then $\vec{\nabla}\bullet\left(\mathscr{L}_{,\vec{\nabla}\{R\}}\{f\}\right)=0$ (that is, the system of equations admits a conservation law),*

$$L=\mathscr{L}+\star$$

and conversely.

Let us apply the above results to the Lagrangean (3.28) for alpha systems. We have

$$\begin{aligned}\mathscr{L}&=\tfrac{1}{2}\{R\}'\vec{\alpha}\bullet\vec{\nabla}\{R\}-H(R,\boldsymbol{x}),\\ \{E|\mathscr{L}\}_{\{R\}}&=\vec{\nabla}\{R\}'\bullet\vec{\alpha}+H_{,\{R\}},\end{aligned} \tag{6.70}$$

so that $\mathscr{L}_{,\vec{\nabla}\{R\}}=\frac{1}{2}\{R\}'\vec{\alpha}$, and hence equation (6.69) reads

$$\vec{\nabla}\bullet\left(\{R\}'\vec{\alpha}\{f\}\right)=0.$$

Expanding yields

$$\vec{\nabla}\{R\}'\bullet\vec{\alpha}\{f\}+\{R\}'\vec{\alpha}\bullet\vec{\nabla}\{f\}=0,$$

which upon using the second of equations (6.70) gives

$$-H_{,\{R\}}\{f\}+\{R\}'\vec{\alpha}\bullet\vec{V}\{f\}=0.$$

Theorem 6.6. *The Lagrangean for alpha systems is invariant under the transformation* (6.62, 63) *to within terms of order ε^2, if and only if $\{f\}$ is such that*

$$\begin{aligned}&\{R\}'\vec{\alpha}\bullet\vec{V}\{f\}=H_{,\{R\}}\{f\},\\&\vec{V}\{R\}'\bullet\vec{\alpha}+H_{,\{R\}}=0\end{aligned}\tag{6.71}$$

are satisfied, in which case $\vec{V}\bullet(\{R\}'\vec{\alpha}\{f\})=0$, and conversely.

We have shown in Theorem 3.1 that the Euler-Lagrange equations are invariant under all transformations T of class $C^{[2]}$. Thus, from the second of equations (6.70), we have under T

$$\vec{V}[\{R\}'+\varepsilon\{f\}'+\star]\bullet\vec{\alpha}+H_{,\{\tilde{R}\}}=\{0\}$$

but

$$H_{,\{\tilde{R}\}}=H_{,\{R\}}+\varepsilon\{H_{,R}\}_{,\{R\}}\{f\}+\star,$$

so that, upon taking the transpose,

$$\vec{\alpha}\bullet\vec{V}\{R\}+\varepsilon\vec{\alpha}\bullet\vec{V}\{f\}-\{H_{,R}\}-\varepsilon\{H_{,R}\}_{,\{R\}}+\star=\{0\}.$$

Noting that $\vec{\alpha}\bullet\vec{V}\{R\}-\{H_{,R}\}=\{0\}$ since $\{R\}$ is such that $\delta\int\mathscr{L}\,d v_n=0$, we obtain

$$\vec{\alpha}\bullet\vec{V}\{f\}-\{H_{,R}\}_{,\{R\}}\{f\}+\star=\{0\},\tag{6.72}$$

which are the **variational equations** associated with an alpha system. Substituting this result into equation (6.30) gives

$$[\{R\}'\{H_{,R}\}_{,\{R\}}-H_{,\{R\}}]\{f\}=0.$$

Hence, we may conclude that equation (6.69) holds when $\{f\}$ is any solution to the variational equations (equations (6.72)) if and only if $H_{,\{R\}}$ is a homogeneous function of degree unity in $\{R\}$.

We have, on comparing equations (6.23) and (6.62, 63), that $\mathscr{L}$ is invariant under (6.23) if and only if

$$\begin{aligned}0&=\vec{V}\bullet[\{R\}'\vec{I}(\vec{I}_l^{-1}\bullet\{\vec{W}_{,R}\})]\\&=\vec{V}\{R\}'\bullet\{\vec{W}_{,R}\}+\{R\}'(\{\vec{W}_{,R}\}_{,\{R\}}\bullet\vec{V}\{R\}+\{U_{,R}\})\end{aligned}\tag{6.73}$$

by equation (6.25).

Functional Changes Under Infinitesimal Canonical Maps

Let $f(p,q)$ be denoted by $f(R)$, and similarly let $f(P,Q)$ be denoted by $f(\tilde{R})$. The question arises as to the increment in $f(R)$ when its arguments are subject to an infinitesimal canonical map. The quantities $\{R\}$ and $\{\tilde{R}\}$ are connected by equation (6.24) so that $f(\tilde{R})=f(R+\Delta R)$. Expanding in a Taylor series about R gives

$$f(\tilde{R})=f(R)+f_{,\{R\}}\{\Delta R\}+\star.$$

Ignoring terms of order ε^2 and substituting from equation (6.24) results in

$$\begin{aligned} f(\widetilde{R}) &= f(R) + f_{,\{R\}}\, \varepsilon\, \vec{I}_l^{-1} \bullet \{\vec{W}_{,R}\} \\ &= f(R) + \mathscr{L}(\bullet \vec{W}, f)\, \varepsilon, \end{aligned}$$

and thus

$$f(\widetilde{R}) - f(R) = \varepsilon\, \mathscr{L}(\bullet \vec{W}, f). \tag{6.74}$$

Similarly if $\vec{f}(R)$ is vectrix function of (p, q),

$$\begin{aligned} \vec{f}(\widetilde{R}) &= \vec{f}(R) + \vec{f}_{,\{R\}}\{\Delta R\}, \\ &= \vec{f}(R) + \vec{f}_{,\{R\}}\big(\vec{I}_l^{-1} \bullet \{\vec{W}_{,R}\}\big)\, \varepsilon, \\ &= \vec{f}(R) + \vec{\mathscr{L}}(\bullet \vec{W}, \vec{f})\, \varepsilon, \end{aligned}$$

and thus

$$\vec{f}(\widetilde{R}) - \vec{f}(R) = \varepsilon\, \vec{\mathscr{L}}(\bullet \vec{W}, \vec{f}) \tag{6.75}$$

where

$$\begin{aligned} \mathscr{L}(\bullet \vec{A}, f) &= f_{,\{R\}}\, \vec{I}_l^{-1} \bullet \{\vec{A}_{,R}\}, \\ \vec{\mathscr{L}}(\bullet \vec{A}, \vec{B}) &= \vec{B}_{,\{R\}}\big(\vec{I}_l^{-1}\{\vec{A}_{,R}\}\big). \end{aligned}$$

Hence, the left-hand bracket vector is said to be the **symbol** for the collection of all infinitesimal canonical maps which arise from equations (6.20) and (6.21).

Section VII. Conservation Laws and Space Structure

The results of this section are considered of such importance that, although they logically could have been included in previous sections, they have been assigned a separate section.

The Basic Conservation Equations

Theorem 7.1. *Let $\{\psi\}$ form an alpha system base $H \in \mathfrak{H} \ominus \mathfrak{H}_0$ with structure vectrix $\vec{\boldsymbol{\alpha}}$, for all $\boldsymbol{x}$ in a given $\mathscr{D}_n^*$, and define the bivectrix $\vec{\vec{W}}$, with components W_{ij}, by*

$$W_{ij} = \tfrac{1}{2}\{\psi\}' \boldsymbol{\alpha}_i \{\psi\}_{,j} - \delta_{ij}\big(\tfrac{1}{2}\{\psi\}' \boldsymbol{\alpha}_k \{\psi\}_{,k} - H\big); \tag{7.1}$$

then $\{\psi\}$ is such that it satisfies identically the n equations

$$W_{ij,i} = \partial_j H \tag{7.2}$$

at all points $\boldsymbol{x}$ in $\mathscr{D}_n^*$.

Proof. From equation (7.1), we obtain by direct operation

$$W_{ij,i} = \tfrac{1}{2}\{\psi\}'_{,i} \boldsymbol{\alpha}_i \{\psi\}_{,j} - \tfrac{1}{2}\{\psi\}_{,j} \boldsymbol{\alpha}_i \{\psi\}_{,i} + H_{,\{\psi\}}\{\psi\}_{,j} + \partial_j H.$$

By hypothesis $\{\psi\}$ forms an alpha system base $H \in \mathfrak{H} \ominus \mathfrak{H}_0$ with structure vectrix $\vec{\boldsymbol{\alpha}}$ at all points in $\mathscr{D}_n^*$, and hence $\{\psi\}$ satisfies the equations

$$\vec{\nabla}\{\psi\}' \bullet \vec{\boldsymbol{\alpha}} + H_{,\{\psi\}} = \{\psi\}'_{,i} \boldsymbol{\alpha}_i + H_{,\{\psi\}} = [0],$$

so that

$$W_{ij,i} = \tfrac{1}{2}\{\psi\}'_{,i} \boldsymbol{\alpha}_i \{\psi\}_{,j} - \tfrac{1}{2}\{\psi\}_{,j} \boldsymbol{\alpha}_i \{\psi\}_{,i} + \partial_j H - \{\psi\}'_{,i} \boldsymbol{\alpha}_i \{\psi\}_{,j}.$$

Now $$\tfrac{1}{2}\{\psi\}'_{,i}\boldsymbol{\alpha}_i\{\psi\}_{,j}-\tfrac{1}{2}\{\psi\}'_{,j}\boldsymbol{\alpha}_i\{\psi\}_{,i}=\tfrac{1}{2}\{\psi\}'_{,i}(\boldsymbol{\alpha}_i-\boldsymbol{\alpha}'_i)\{\psi\}_{,j};$$

but $\boldsymbol{\alpha}_i=-\boldsymbol{\alpha}'_i$, since $\vec{\boldsymbol{\alpha}}$ is a structure vectrix, and hence

$$\tfrac{1}{2}\{\psi\}'_{,i}\boldsymbol{\alpha}_i\{\psi\}_{,j}-\tfrac{1}{2}\{\psi\}'_{,j}\boldsymbol{\alpha}_i\{\psi\}_{,i}=\{\psi\}'_{,i}\boldsymbol{\alpha}_i\{\psi\}_{,j}.$$

Substituting this result into the last equation above for $W_{ij,i}$ gives

$$W_{ij,i}=\partial_j H. \qquad \text{Q.E.D.}$$

The divergence type equations (7.2) constitute the basic conservation laws for alpha systems. Considerations similar to those above have been used by many authors for a variety of purposes. To note an instance of particular importance, EINSTEIN[19] developed the momentum-energy pseudo-tensor for the gravitational field and the matter field by similar considerations.

It is to be noted that equations (7.2) result without any assumptions as to the differential geometry of the space of independent variables. In particular we have not even assumed that the space of independent variables is metrizable except in a local sense but only that it satisfies the postulates of a Hausdorff space as stated in Section I.

The Principle of Extremal Invariance

We have shown in Theorem 3.7 that alpha systems arise from the extremization of integrals of the form

$$\mathscr{I}=\int_{\mathscr{D}_n^*}\mathscr{L}\,dv_n, \tag{7.3}$$

where

$$\mathscr{L}=\tfrac{1}{2}\{\psi\}'\vec{\boldsymbol{\alpha}}\bullet\vec{V}\{\psi\}-H(\psi,\boldsymbol{x}); \tag{7.4}$$

the variation of $\{\psi\}$ vanishes on the boundary of $\mathscr{D}_n^*$, and no variation of the independent variables is considered. We now consider more general variation processes in which the variation of $\{\psi\}$ does not necessarily vanish on the boundary and the independent variables are also allowed to vary.

Set

$$\{\tilde{\psi}(\boldsymbol{x},\varepsilon)\}=\{\psi(\boldsymbol{x})\}+\varepsilon\{\eta(\boldsymbol{x})\}, \tag{7.5}$$

$$\tilde{\boldsymbol{x}}=\boldsymbol{x}+\varepsilon\,\delta\boldsymbol{x}(\boldsymbol{x}), \tag{7.6}$$

$$\{\tilde{\psi}(\tilde{\boldsymbol{x}})\}=\{\psi(\boldsymbol{x})\}+\varepsilon\,\delta\{\psi(\boldsymbol{x})\}, \tag{7.7}$$

which are related by

$$\delta\{\psi(\boldsymbol{x})\}=\{\eta(\boldsymbol{x})\}+\vec{V}\{\psi(\boldsymbol{x})\}\otimes\delta\boldsymbol{x}(\boldsymbol{x}) \tag{7.8}$$

to within terms of order ε^2. Computing the variation of $\mathscr{I}$ resulting from equations (7.5) and (7.6) gives

$$\delta\mathscr{I}=\int_{\mathscr{D}_n^*}(\mathscr{L}_{,\{\psi\}}\{\eta\}+\mathscr{L}_{,\vec{V}\{\psi\}}\bullet\vec{V}\{\eta\})\,dv_n+\int_{\mathscr{D}_n^*\ominus\mathscr{D}_n}\mathscr{L}\,\delta\boldsymbol{x}\otimes\vec{N}\,dS,$$

[19] EINSTEIN, A.: Annalen d. Physik 49 (1916). — Sitzungsber. d. Preuß. Akad. d. Wiss. 1916.

where equation (1.16) has been used. Since

$$\mathscr{L}_{,\vec{\nabla}\{\psi\}} \bullet \vec{\nabla}\{\eta\} = \vec{\nabla}\left(\mathscr{L}_{,\vec{\nabla}\{\psi\}}\{\eta\}\right) - (\vec{\nabla} \bullet \mathscr{L}_{,\vec{\nabla}\{\psi\}})\{\eta\},$$

we have the equivalent form, upon applying equations (1.14) and (1.15),

$$\delta\mathscr{I} = \int_{\mathscr{D}_n^*} (\mathscr{L}_{\{\psi\}} - \vec{\nabla} \bullet \mathscr{L}_{,\vec{\nabla}\{\psi\}})\{\eta\}\, dv_n - \int_{\mathscr{D}_n^* \ominus \mathscr{D}_n} (\mathscr{L}\,\delta \boldsymbol{x} + \mathscr{L}_{,\vec{\nabla}\{\psi\}}\{\eta\}) \otimes \vec{N}\, dS. \quad (7.9)$$

Solving equation (7.8) for $\{\eta\}$ and substituting into the second integral on the right for equation (7.9) yields

$$\begin{aligned} \delta\mathscr{I} = &\int_{\mathscr{D}_n^*} \{E|\mathscr{L}\}_{\{\psi\}}\{\eta\}\, dv_n + \\ &+ \int_{\mathscr{D}_n^* \ominus \mathscr{D}_n} \{\mathscr{L}\,\delta\boldsymbol{x} + \mathscr{L}_{,\vec{\nabla}\{\psi\}}(\delta\{\psi\} - \vec{\nabla}\{\psi\} \otimes \delta\boldsymbol{x})\} \otimes \vec{N}\, dS \end{aligned}$$

upon noting that $\{E|\mathscr{L}\}_{\{\psi\}} = \mathscr{L}_{\{\psi\}} - \vec{\nabla} \bullet \mathscr{L}_{,\vec{\nabla}\{\psi\}}$. Expanding this form in terms of components then yields

$$\begin{aligned} \delta\mathscr{I} = &\int_{\mathscr{D}_n^*} \{E|\mathscr{L}\}_{\{\psi\}}\{\eta\}\, dv_n + \\ &- \int_{\mathscr{D}_n^* \ominus \mathscr{D}_n} \{(\mathscr{L}\,\delta_{ik} - \mathscr{L}_{,\{\psi\},i}\{\psi\}_{,k})\,\delta x^k - \mathscr{L}_{,\{\psi\},i}\,\delta\{\psi\}\} N_i\, dS. \end{aligned} \quad (7.10)$$

Set

$$\{\vec{Z}\}' = \mathscr{L}_{,\vec{\nabla}\{\psi\}} \quad (7.11)$$

and

$$T_{ik} = \mathscr{L}\,\delta_{ik} - \{Z_i\}'\{\psi\}_{,k}; \quad (7.12)$$

then

$$\delta\mathscr{I} = \int_{\mathscr{D}_n^*} \{E|\mathscr{L}\}_{\{\psi\}}\{\eta\}\, dv_n - \int_{\mathscr{D}_n^* \ominus \mathscr{D}_n} \{T_{ik}\,\delta x^k - \{Z_i\}'\,\delta\{\psi\}\} N_i\, dS. \quad (7.13)$$

Making the additional substitutions

$$T_{ik} N_i\, dS = dY_k, \quad \{Z_i\}' N_i\, dS = d\{\mathscr{P}\}', \quad (7.14)$$

we obtain the following form of the variation of $\mathscr{I}$:

$$\delta\mathscr{I} = \int_{\mathscr{D}_n^*} \{E|\mathscr{L}\}_{\{\psi\}}\{\eta\}\, dv_n - \int_{\mathscr{D}_n^* \ominus \mathscr{D}_n} (dY_k\,\delta x^k - d\{\mathscr{P}\}'\,\delta\{\psi\}). \quad (7.15)$$

Equation (7.15) states that

$$(Y_k, x^k) \quad \text{and} \quad (\{\mathscr{P}\}, \{\psi\})$$

may be considered as conjugate quantities on $\mathscr{D}_n^* \ominus \mathscr{D}_n$ in the same sense that momentum and coordinates are conjugate variables in the Hamiltonian mechanics of particles $(n = 1)$[20].

There are several significant conclusions which may be drawn from equation (7.15). Under the conditions

$$\{E|\mathscr{L}\}_{\{\psi\}} = [0], \quad \delta\{\psi\}|_{\mathscr{D}_n^* \ominus \mathscr{D}_n} = \{0\}, \quad (7.16)$$

[20] Weiss, P.: Proc. Roy. Soc. 169, 102—199 (1938).

(*i.e.*, $\{\psi\}$ forms an alpha system and the total variation of $\{\psi\}$ vanishes on the boundary of $\mathscr{D}_n^*$), the vanishing of $\delta\mathscr{I}$ implies

$$0 = \int_{\mathscr{D}_n^* \bigcirc \mathscr{D}_n} dY_k\,\delta x^k = \int_{\mathscr{D}_n^* \ominus \mathscr{D}_n} T_{ik} N_i\,\delta x^k\,dS \tag{7.17}$$

upon making use of equation (7.14). Transforming equation (7.17) to an integral over $\mathscr{D}_n^*$ by use of equation (1.16) gives

$$0 = \int_{\mathscr{D}_n^*} (T_{ik}\,\delta x^k)_{,i}\,dv_n = \int_{\mathscr{D}_n^*} (T_{ik,i}\,\delta x^k + T_{ik}\,\delta x^k_{,i})\,dv_n.$$

The last equation can be satisfied for general $\mathscr{D}_n^*$ only if

$$T_{ik,i}\,\delta x^k = -\,T_{ik}\,\delta x^k_{,i} \tag{7.18}$$

holds at all points of $\mathscr{D}_n^*$.

Theorem 7.2. *We have, for alpha systems with structure vectrix* $\vec{\boldsymbol{\alpha}}$,

$$W_{ik} = -\,T_{ik}, \tag{7.19}$$

$$\{\vec{Z}\}' = \tfrac{1}{2}\{\psi\}'\,\vec{\boldsymbol{\alpha}}, \tag{7.20}$$

$$d\{\mathscr{P}\}' = \tfrac{1}{2}\{\psi\}'\,\vec{\boldsymbol{\alpha}}\bullet\vec{N}\,dS, \tag{7.21}$$

$$dY_k = -\,W_{ik} N_i\,dS. \tag{7.22}$$

Proof. The results are immediate upon substituting equation (7.4), which defines the Lagrangean for alpha systems, into equations (7.11), (7.12), and (7.14) upon using equation (7.1). Q.E.D.

Theorem 7.2, in conjunction with equation (7.18), admits an important and far-reaching conclusion. Substituting equation (7.2) and (7.19) into equation (7.18) gives

$$W_{ik}\,\delta x^k_{,i} = -\,(\partial_k H)\,\delta x^k.$$

We have thus proved

Theorem 7.3. *Let* $\{\psi\}$ *form an alpha system with structure vectrix* $\vec{\boldsymbol{\alpha}}$ *and base function* H, *and set*

$$\mathscr{I} = \int_{\mathscr{D}_n^*} \mathscr{L}\,dv_n$$

where

$$\mathscr{L} = \tfrac{1}{2}\{\psi\}'\,\vec{\boldsymbol{\alpha}}\bullet\vec{V}\{\psi\} - H(\psi, \boldsymbol{x});$$

then the variation of $\mathscr{I}$, *as induced by equations* (7.5), (7.6) *and* (7.7), *vanishes for* $\delta\{\psi\} = \{0\}$ *on* $\mathscr{D}_n^* \ominus \mathscr{D}_n$ *if and only if* $\delta\boldsymbol{x}$ *satisfies*

$$W_{ik}\,\delta x^k_{,i} = -\,(\partial_k H)\,\delta x^k \tag{7.23}$$

where W_{ik} *is defined by equation* (7.1).

Examining equation (7.15) in detail, we see that requiring $\mathscr{I}$ to be extremal under the condition $\delta\{\psi\} = \{0\}$ on $\mathscr{D}_n^* \ominus \mathscr{D}_n$, for $\delta\boldsymbol{x} = \boldsymbol{O}$, results in exactly the variation process used in Section III requiring $\{\psi\}$ to form an alpha system,

(*i.e.*, for $\delta\boldsymbol{x}=\boldsymbol{O}$ we have, by equation (7.8), $\delta\{\psi\}=\{\eta\}$). We assumed in Theorem 7.3 that $\{\psi\}$ formed an alpha system and thereby obtained equation (7.23) under the hypothesis that the total variation of $\{\psi\}$ vanished on $\mathscr{D}_n^*\ominus\mathscr{D}_n$. Because of the importance of the content of Theorem 7.3, we state the following principle whereby we are assured that $\{\psi\}$ will always form an alpha system under the variation (7.5) and (7.6).

The Principle of Extremal Invariance: *The quantity $\mathscr{I}$, defined by*

$$\mathscr{I}=\int_{\mathscr{D}_n^*}\mathscr{L}\,dv_n,$$

where

$$\mathscr{L}=\tfrac{1}{2}\{\psi\}'\,\vec{\boldsymbol{\alpha}}\bullet\vec{V}\{\psi\}-H(\psi,\boldsymbol{x}),$$

shall remain extremal for all variations (7.5) *and* (7.6) *such that* (1) *the total variation, $\delta\{\psi\}$, of $\{\psi\}$ vanishes on $\mathscr{D}_n^*\ominus\mathscr{D}_n$ and* (2) *equation* (7.23) *is satisfied.*

Under the principle of extremal invariance we have, by equation (7.15), that $\{\psi\}$ satisfies

$$\vec{V}\{\psi\}'\bullet\vec{\boldsymbol{\alpha}}+H_{,\{\psi\}}=[0] \tag{7.24}$$

and hence forms an alpha system.

Given a solution, say $\{\psi_0(\boldsymbol{x})\}$, of equations (7.24), we may use equation (7.23) to determine sets of functions $\delta x_0^k(\boldsymbol{x})$. These functions may be interpreted as elements of a group of infinitesimal transformations of the underlying space $\mathscr{E}_n$. Under the principle of extremal invariance we see that such a group of transformations preserves equations (7.24), and conversely, preserves the extremal character of $\mathscr{I}$ by Theorem 7.3. Since the W_{ij} and the H appearing in equation (7.23) depend on the $\{\psi\}$ and hence on the alpha system considered, we may conclude further that the class of infinitesimal transformations which $\mathscr{E}_n$ can admit under the principle of extremal invariance is determined by the alpha system defined over $\mathscr{E}_n$. We have thus proved

Theorem 7.4. *The class of infinitesimal transformations which $\mathscr{E}_n$ can admit under the principle of extremal invariance is determined by the alpha system defined over $\mathscr{E}_n$, in the sense of equation* (7.23).

We made detailed note in Section I that we do not assume a differential geometry for $\mathscr{E}_n$ but only require $\mathscr{E}_n$ to be a Hausdorff space. By Theorem 7.4, we see that the introduction of the Principle of Extremal Invariance restricts the class of infinitesimal transformations which $\mathscr{E}_n$ can admit in terms of the alpha system defined over $\mathscr{E}_n$. Since the infinitesimal transformations admitted by a space determine the transformation groups which a space admits, and the transformation groups in turn determine the nature and type of differential geometry which may be introduced into a space, we see that **the alpha system defined over $\mathscr{E}_n$ determines a natural type of differential geometry which may be introduced into that $\mathscr{E}_n$ under the principle of extremal invariance.** In this sense, equation (7.23) is analogous to the Killing equations[21]. This is particularly evident if H does not depend explicitly on $\boldsymbol{x}$.

---

[21] Killing, W.: J. für die Reine und Angew. Math. **109**, 121–186 (1892).

Section VIII. Integrability Conditions and Related Topics

The discussion in Section II, following the first definition of a canonical map, pointed out that one need consider only those bases functions for which there exist (p, q) which satisfy equations (2.1) or their matrix form, equation (2.12). We take up in this section the question of the form that a base function may take in order that there exist p's and q's such that they satisfy equations (2.12).

Decomposition of Hamiltonian Systems

If (p, q) form a Hamiltonian system base H, then by equations (2.12)

$$\vec{\boldsymbol{I}} \bullet \vec{V}\{R\} = -\{H_{,R}\}. \tag{8.1}$$

Applying Theorem 1.2, we may solve these equations for $\vec{V}\{R\}$, giving

$$\vec{V}\{R\} = -\vec{\boldsymbol{I}}_r^{-1}\{H_{,R}\} + \{\vec{C}\} \tag{8.2}$$

where $\{\vec{C}\}$ is an arbitrary column vectrix which satisfies the condition $\vec{\boldsymbol{I}} \bullet \{\vec{C}\} = \{0\}$. Equation (8.2) is the most general explicit decomposition of Hamiltonian systems into systems in which each partial derivative of each dependent variable appears separately.

Let $\{R_1\}$ be any solution to equation (8.1); then we can solve for a $\{\vec{C}\}$ by equation (8.2). On the other hand, let $\{R_2\}$ be a solution of equation (8.2) for a particular $\{\vec{C}\}$; then $\{R_2\}$ satisfies equation (8.1), as is evident by inner multiplication of equation (8.2) by $\vec{\boldsymbol{I}}$ from the left. Thus, every solution of equation (8.2) is a solution of equation (8.1), and for every solution of equation (8.1) we can solve for $\{\vec{C}\}$ so that it is a solution of equation (8.2). In this sense equations (8.1) and (8.2) may be considered similar relative to their collections of solutions. More important, we see that *if we prove the existence of solutions of equation* (8.2) *for any one* $\{\vec{C}\}$, *then we have proved the existence of solutions of equation* (8.1).

The $\{\vec{C}\}$ appearing in equation (8.2) is subject only to the conditions $\vec{\boldsymbol{I}} \bullet \{\vec{C}\} = \{0\}$. Using the defined form of $\vec{\boldsymbol{I}}$ and writing

$$\{\vec{C}\} = \vec{e}_i\{C_i\}, \qquad \{C_i\} = \begin{Bmatrix} \{C_{\alpha 0 i}\} \\ \{C_{\alpha 1 i}\} \\ \vdots \\ \{C_{\alpha n i}\} \end{Bmatrix},$$

we have $\vec{\boldsymbol{I}} \bullet \{\vec{C}\} = \{\vec{0}\}$ if and only if the conditions

$$C_{\alpha 0 i} = 0, \qquad C_{\alpha i i} - 0 \tag{8.3}$$

hold. Substituting this result into equation (8.2) gives

$$\vec{V}\begin{Bmatrix} \{q_\alpha\} \\ \{p_{\alpha 1}\} \\ \vdots \\ \{p_{\alpha n}\} \end{Bmatrix} = \vec{\boldsymbol{I}}_r^{-1}\{H_{,R}\} + \vec{e}_i \begin{Bmatrix} \{0\} \\ \{C_{\alpha 1 i}\} \\ \vdots \\ \{C_{\alpha n i}\} \end{Bmatrix},$$

from which we see that $\vec{V}\{q_\alpha\}$ is determined uniquely, while $\vec{V}\{p_{\alpha k}\}$ are determined only to within the arbitrary functions $\vec{e}_i\{C_{\alpha k i}\}$ which satisfy equations (8.3). This situation is not a surprising one since the $p_{\alpha i}$ satisfy the divergence type equations

$$p_{\alpha i, i} = -H_{,q_\alpha}$$

if the system is Hamiltonian.

Expanding equation (8.2) in terms of the components of the various vectrices appearing gives

$$\{R\}_{,i} = -\boldsymbol{I}_{ri}^{-1}\{H_{,R}\} + \{C_i\}$$

where

$$\vec{\boldsymbol{I}}_r^{-1} = \vec{e}_i \boldsymbol{I}_{ri}^{-1}, \qquad \{\vec{C}\} = \vec{e}_i\{C_i\}.$$

Let $F(R, \boldsymbol{x})$ be any given function of R and $\boldsymbol{x}$ of class C^1; then

$$\begin{aligned} F_{,i} &= F_{,\{R\}}\{R\}_{,i} + \partial_i F \\ &= -F_{,\{R\}}\boldsymbol{I}_{ri}^{-1}\{H_{,R}\} + F_{,\{R\}}\{C_i\} + \partial_i F \\ &= -\mathscr{R}_i(H, F) + F_{,\{R\}}\{C_i\} + \partial_i F \end{aligned}$$

by equation (4.22), $(\mathscr{R}_i(H, F)$ is the i^{th} component of the right-hand bracket vectrix formed from $\boldsymbol{I}_r^{-1}$ and $\mathscr{R}_i(H, F) = -\mathscr{L}_i'(F, H))$. Thus we have proved the following

***Theorem 8.1.** *Let $F(R, \boldsymbol{x})$ be an arbitrary given function of $\boldsymbol{x}$ and $\{R\}$ of class C^1, and let $\{R\}$ form a Hamiltonian system base H; then*

$$F_{,i} = \mathscr{L}_i'(F, H) + \{C_i\}'\{F_{,R}\} + \partial_i F \tag{8.4}$$

where

$$\boldsymbol{I}_i\{C_i\} = \{0\}.$$

The Classic Integrability Theorem

We now turn to the explicit problem of integrability of equations (8.1). From their equivalent form, as given by equations (8.2), we have reduced the problem of integrability to one taken up in the literature, particularly by EISENHART[22]. For purposes of clarity, we include a brief synopsis of the principal results.

Consider a system of partial differential equations

(a) $$\frac{\partial \mu^\alpha}{\partial x^i} = \psi_i^\alpha(\mu, x), \qquad \alpha = 1, \ldots, m, \quad i = 1, \ldots, n,$$

where the ψ_i^α are analytic functions of the μ^α and the x^i. These equations are equivalent to the system of total differential equations

(b) $$d\mu^\alpha = \psi_i^\alpha\, dx^i.$$

The conditions of integrability of equations (a) considered in terms of equations (b) are

(c) $$\frac{\partial \psi_i^\alpha}{\partial x^j} + \frac{\partial \psi_i^\alpha}{\partial \mu^\beta}\psi_j^\beta = \frac{\partial \psi_j^\alpha}{\partial x^i} + \frac{\partial \psi_j^\alpha}{\partial \mu^\gamma}\psi_i^\gamma.$$

[22] EISENHART, L. P.: Non-Riemannian Geometry. Amer. Math. Soc. Colloquium Publ. No. 8.

If these equations are satisfied identically, the system (a) is said to be **completely integrable**. In this case, the solution is expressible in the form

$$\text{(d)} \qquad \mu^\alpha = C^\alpha + \left(\frac{\partial \mu^\alpha}{\partial x^i}\right)\bigg|_0 (x^i - x_0^i) + \cdots,$$

where $\left(\frac{\partial \mu^\alpha}{\partial x^i}\right)\Big|_0 = \psi_i^\alpha(C, x)$ and the other coefficients are obtained by differentiation of equation (a) and evaluating by equating the x's to the x_0's and the μ's to the C's. Thus, for the domain of the x's in which the series (d) converges, we have a solution of equations (a) determined by m constants. We denote such a solution by

$$\text{(e)} \qquad \mu^\alpha = \mu^\alpha(x^1, x^2, \ldots, x^n; C^1, \ldots, C^m).$$

If equations (c) are not satisfied identically, we have a set F_1 of equations which establish conditions upon the μ's as functions of the x's. If we differentiate each of these equations with respect to the x's and substitute for $\partial\mu^\alpha/\partial x^i$ from equations (a), either the resulting equations are a consequence of the set F_1, or we get a new set F_2. Proceeding in this way, we get a sequence of sets of equations $F_1, F_2, F_3, \ldots$, which must be algebraically compatible if equations (a) are to admit a solution. If one of these sets is not a consequence of the preceding sets, it introduces at least one additional condition. Consequently, if the equations (a) are to admit a solution, there must be a positive number U such that the equations of the $(U+1)^{\text{th}}$ set are satisfied identically because of the equations of the preceding U sets; otherwise we should obtain more than m independent equations which would imply a relation between the x's. Moreover, from this argument it follows that $U \leqq m$.

Conversely, suppose that there is a number U such that the equations of the sets

$$\text{(f)} \qquad F_1,\ F_2,\ F_3, \ldots$$

are compatible and each set introduces one or more conditions independent of the conditions imposed by the equations of the other sets, and that all of the equations of the set

$$\text{(g)} \qquad F_{U+1}$$

are satisfied identically because of the equations of the sets (f). Assume that there are p $(<m)$ independent conditions imposed by (f), say

$$G_\gamma(\mu, x) = 0, \qquad \gamma = 1, \ldots, p.$$

Since the Jacobian matrix $(\partial G_\gamma/\partial\mu^\alpha)$ is of rank p, these equations may be solved for p of the μ's in terms of the remaining μ's and the x's, which, by suitable ordering, may be written as

$$\text{(h)} \qquad \mu^\sigma - \varphi^\sigma(\mu^{p+1}, \ldots, \mu^m; x) = 0, \qquad \sigma = 1, \ldots, p.$$

From these equations we have by differentiation

$$\frac{\partial \mu^\sigma}{\partial x^i} - \frac{\partial \varphi^\sigma}{\partial x^\nu}\,\frac{\partial \mu^\nu}{\partial x^i} - \frac{\partial \varphi^\sigma}{\partial x^i} = 0, \qquad \nu = p+1, \ldots, m.$$

Replacing $\partial\mu^\sigma/\partial x^i$ by means of equations (a) gives

$$\text{(i)} \qquad \psi_i^\sigma - \frac{\partial\varphi^\sigma}{\partial\mu^\nu}\psi_i^\nu - \frac{\partial\varphi^\sigma}{\partial x^i} = 0, \qquad \sigma=1,\ldots,p, \quad \nu=p+1,\ldots,m$$

which are satisfied because of the sets (f) and (g), as follows from the method of obtaining the latter. Accordingly, we have by subtraction

$$\text{(j)} \quad \frac{\partial\mu^\sigma}{\partial x^i} - \psi_i^\sigma - \frac{\partial\varphi^\sigma}{\partial\mu^\nu}\left(\frac{\partial\mu^\nu}{\partial x^i} - \psi_i^\nu\right) = 0, \qquad \sigma=1,\ldots,p, \quad \nu=p+1,\ldots,m.$$

From these equations it follows that if the functions $\mu^{p+1},\ldots,\mu^m$ are chosen to satisfy the equations

$$\text{(k)} \qquad \frac{\partial\mu^\nu}{\partial x^i} = \tilde{\psi}_\nu^i\,(\mu^{p+1},\ldots,\mu^m; x), \quad \nu=p+1,\ldots,m,$$

where the $\tilde{\psi}_i^\nu$ are obtained from the ψ_i^ν on replacing μ^σ $(\sigma=1,\ldots,p)$ by their expressions (h), then equations (a) for $\alpha=1,\ldots,p$ are satisfied by the values given by equations (h). Since the equations of the set F_1 are satisfied identically because of equation (h), it follows that equations (k) are completely integrable, for the equations arising from expressing their conditions of integrability are in the set because of equations (i). Consequently, there is a solution in this case, and it involves $m-p$ arbitrary constants. When $p=m$, we have in place of (h) $\mu^\alpha=\varphi^\alpha(x)$, and in place of equations (j) the functions μ^α satisfying equations (a). Thus we have

Theorem A. *In order that a system of equations* (a) *admit a solution, it is necessary and sufficient that there exist a positive integer U ($\leq m$) such that the equations $F_1, F_2, \ldots, F_U$ are compatible for all values of the x's in a domain, and that the equations of the set F_{U+1} are satisfied identically because of the former sets; if p is the number of independent equations in the first U sets, the solution involves $m-p$ arbitrary constants.*

Integrability and Generalized Hamiltonian Systems

Since equations (8.2) are of the form of equations (a) in the above discussion, and we have seen that if we demonstrate the existence of solutions to equations (8.2) for at least one $\{\vec{C}\}$, then we insure the existence of solutions of equations (8.1). We may apply the results of the above theorem directly.

Definition. Equations (8.1) are said to be completely integrable if there exists at least one column vectrix $\{\vec{C}\}$, satisfying the condition $\vec{I}\bullet\{\vec{C}\}=0$, such that equations (8.2) are completely integrable.

Note. Since different solutions of equations (8.1) may correspond to different choices of $\{\vec{C}\}$, complete integrability in the sense of the above definition does not imply that all solutions of equations (8.1) constitute an S-parameter family of solutions where S is finite. However, to every choice of $\{\vec{C}\}$ there does exist an $N(n+1)$-parameter sub-family of solutions to equations (8.1) if these equations are completely integrable.

***Theorem 8.2.** *Equations* (8.1) *are completely integrable for a given* $H \in \mathfrak{H}$ *if and only if there exists at least one column vectrix* $\{\vec{C}\}$ *such that*

$$\vec{I} \bullet \{\vec{C}\} = \{0\}$$

and

$$\begin{aligned} \mathscr{L}'_j(\mathscr{L}'_i(\{R\}, H), H) - \mathscr{L}'_i(\mathscr{L}'_j(\{R\}, H), H) + \\ + \mathscr{L}'_j(\{C_i\}, H) - \mathscr{L}'_i(\{C_j\}, H) + \\ + \mathscr{L}'_i(\{R\}, H)_{,\{R\}}\{C_j\} - \mathscr{L}'_j(\{R\}, H)_{,\{R\}}\{C_i\} + \\ + \{C_i\}_{,\{R\}}\{C_j\} - \{C_j\}_{,\{R\}}\{C_i\} + \partial_j \mathscr{L}'_i(\{R\}, H) - \\ - \partial_i \mathscr{L}'_j(\{R\}, H) + \partial_j\{C_i\} + \partial_i\{C_j\} = \{0\} \end{aligned} \tag{8.5}$$

is identically satisfied in $\{R\}$ *and* $\boldsymbol{x}$.

Proof. From Theorem 8.1 we have

$$\begin{aligned} \{R\}_{,j} &= -\mathscr{R}_j(H, \{R\}) + \{C_j\}, \\ &= \mathscr{L}'_j(\{R\}, H) + \{C_j\}, \end{aligned}$$

so that the result follows from equation (c) above by use of equation (8.4).

Starting with equations (8.5) as the system F_1 and using Theorem 8.1 to generate the systems $F_2, \ldots$, we have by Theorem A, upon noting that m in that discussion must be replaced by $N(n+1)$,

***Theorem 8.3.** *In order for equations* (8.1) *to admit a solution, it is both necessary and sufficient that there exist a positive integer* $U(\leq N(n+1))$ *such that the bases function* H *and the column vectrix* $\{\vec{C}\}$, *defined by equation* (8.2), *be such that the equations* $F_1, F_2, F_3, \ldots$ *are compatible for all values of the* x's *in the domain under consideration, and that the equations of the set* F_{U+1} *are identically satisfied because of the former sets. If* S *is the number of independent equations of the first* U *sets, then the solution will involve* $N(n+1) - S$ *arbitrary constants.*

This memoir is based on a thesis submitted to the Johns Hopkins University in partial fulfillment of the requirements for the Ph.D. degree.

Hughes Aircraft Company
Systems Development Laboratories
Culver City, California

(Received November 21, 1959)

Scalar Diffraction Theory and Turning-Point Problems

N. D. Kazarinoff & R. K. Ritt

Communicated by L. Cesari

In this paper we wish to discuss and prove a fundamental theorem [*15*] in the theory of scalar diffraction and to motivate and describe a number of turning-point problems in the theory of ordinary differential equations, some of which are as yet unsolved. Let B be a smooth, simple, closed surface in E^3, and suppose that B is steadily illuminated by mono-chromatic electromagnetic (or sound) waves from a source distribution $\varrho(x)$. It is of importance to know the field surrounding B; in particular, one would like to know what percent of the incident energy is returned in a given direction, what percent is scattered in a forward direction, what are the surface currents, and how all these things depend upon B and the wavelength λ of the incident radiation. This is an immensely difficult problem. Heuristic mathematical-physical theories in good agreement with observed phenomena have been found in certain cases, notably by V. A. Fock [*2*] and J. B. Keller [*10*]; but a rigorous mathematical treatment has been given only for special shapes and wavelengths, for example, if λ is small with respect to the dimensions of B and B is a prolate spheroid or elliptic cylinder of small eccentricity, a right circular cone, or a parabolic cylinder.

Fortunately, there is a relation between the electromagnetic problem described above and the acoustical or scalar one. This connection is described in [*10, 18*]. It enables one to solve the electromagnetic problem for a particular body B, or at least to determine part of the solution, by solving the scalar problem. For the remainder of this paper we shall confine our remarks to the scalar problem (I) defined as follows. Let f, g, and ϱ be functions, having compact support, which are defined on the closed exterior of B, henceforth denoted by V. Let f and g be in $C^{(2)}(V)$, and let the normal derivatives $\partial f/\partial n$ and $\partial g/\partial n$ be zero on B. We postulate the existence of a function $u(x,t)$ which is in $C^{(2)}(V)$ for each fixed positive t, which is in $C^{(2)}(t>0)$, and which satisfies the following boundary-initial value problem:

$$\text{(I)}\qquad \begin{aligned} &\nabla^2 u - u_{tt} = \varrho(x)\, e^{i\omega t} && \text{for } t>0;\\ &\partial u/\partial n = 0 && \text{on } B;\\ &u(x,0^+) = f(x), && u_t(x,0^+) = g(x), \end{aligned}$$

the convergence being uniform.

Existence theorems for problem (I) have been given by Ladyženskaya & Višik [7]. They are primarily concerned with the existence of weak solutions. They then show several ways how these weak solutions may be proved to be strong ones when the data are sufficiently regular. The work of Lewis [11] is much more severely restricted — B must be a hyperplane. Of course, the presence of initial conditions prevents separation of t in (I); and at the present time it is unknown whether or not there exists a steady "state" solution of the form

$$u(x,t) = u(x)\, e^{i\omega t} + u^*(x,t),$$

where $u^* \to 0$ as $t \to \infty$. The theorem we shall prove asserts that the $C-1$ limit of u^* is zero, however. This generalizes a result of Ladyženskaya [7; p. 96].

We first observe

Lemma 1. *The solution of* (I) *is unique. If f, g, and ϱ vanish for $|x| > R$, then $u(x,t) = 0$ for $|x| > R + t$.*

The proof of Lemma 1, under less restrictive hypotheses, is contained in [18; p. 135]. We now state the theorem we shall prove.

Theorem. *Let $v(x,t) = e^{-i\omega t} u(x,t)$, where $u(x,t)$ is the solution of* (I). *Then,*

a) $\lim\limits_{T\to\infty} \dfrac{1}{T}\displaystyle\int_0^T v(x,t)\, dt = v^*(x)$ *exists;*

b) $(\nabla^2 + \omega^2)\, v^*(x) = \varrho(x), \qquad \partial v^*/\partial n = 0$ *on* B,

v^* satisfies the Sommerfeld radiation condition;

c) *The Laplace transform, $\hat{v}(x,s)$, of $v(x,t)$ exists for $s > 0$ and lies in $L^{(2)}(V)$;*

d) $\lim\limits_{s\to 0+} s\,\hat{v}(x,s) = v^*(x)$;

e) *In the computation of the limit in* (d), *$s\hat{v}(x,s)$ may be replaced by the function $\Phi(x,s)$ which satisfies the equation*

$$[\nabla^2 + (\omega - i s)^2]\, \Phi(x,s) = \varrho(x) \tag{1}$$

and the conditions that $\partial\Phi/\partial n = 0$ on B and Φ lie in $L^{(2)}(V)$.

Our proof requires several lemmas.

Lemma 2. *If $u(x,t)$ is the solution of* (I) *with $\varrho(x) = 0$ (the homogeneous problem) and if $v(x,t) = u(x,t)\exp(-i\omega t)$, then the $L^{(2)}$ norms of v and v_t are bounded functions of t.*

Proof. The uniqueness theorem (Lemma 1) permits one to express the column vector $\mathrm{col}(u, u_t)$ in the form

$$\begin{pmatrix} u \\ u_t \end{pmatrix} = \mathfrak{J}(t) \begin{pmatrix} f \\ g \end{pmatrix},$$

where $\mathfrak{J}(t)$ is a strongly continuous semigroup on the space $\mathfrak{X} = \mathfrak{X}_0 \times \mathfrak{X}_0$, $\mathfrak{X}_0$ being the space of functions in $C^{(2)}(V)$ having compact support and having normal

derivatives which vanish on B. The topology of $\mathfrak{X}_0$ is that of uniform convergence of all derivatives of order less than or equal to two. By a slight modification of a theorem of HILLE's [3; p. 396], one can express $\mathfrak{J}(t)$ as

$$\mathfrak{J}(t) = \exp(\mathfrak{n}\, t),$$

where $\mathfrak{n}$ is the closed operator

$$\mathfrak{n} = \begin{pmatrix} 0 & 1 \\ \nabla^2 & 0 \end{pmatrix}$$

whose domain is dense in the completion of $\mathfrak{X}$. It is then an elementary computation to show that $\operatorname{col}(v, v_t)$ can be expressed in the form

$$\begin{pmatrix} v \\ v_t \end{pmatrix} = \exp(\mathfrak{m}\, t) \begin{pmatrix} f^* \\ g^* \end{pmatrix},$$

where $\mathfrak{m}$ is described by the matrix

$$\mathfrak{m} = \begin{pmatrix} 0 & 1 \\ \nabla^2 + \omega^2 & -2i\omega \end{pmatrix},$$

and where f^* and g^* are in $\mathfrak{X}_0$.

Now let us regard $\operatorname{col}(f^*, g^*)$ as an element, not of $\mathfrak{X}$, but of $\mathfrak{Y} = \mathfrak{Y}_0 \times \mathfrak{Y}_0$, where $\mathfrak{Y}_0 = L^{(2)}(V)$. We seek the resolvent $(\mathfrak{m} - \lambda)^{-1}$ in $\mathfrak{Y}$. To find this resolvent, we must solve the equation

$$\begin{pmatrix} f_1 \\ g_1 \end{pmatrix} = (\mathfrak{m} - \lambda) \begin{pmatrix} f_2 \\ g_2 \end{pmatrix},$$

which is equivalent to the system

$$g_2 - \lambda f_2 = f_1,$$

$$[\nabla^2 + \omega^2] f_2 - (2i\omega + \lambda) f_2 = (2i\omega + \lambda) f_1 + g_1.$$

If we eliminate g_2 from these equations, we obtain the equation

$$[\nabla^2 + (\omega - i\lambda)^2] f_2 = (2i\omega + \lambda) f_1 + g_1.$$

Now, the operator $-\nabla^2$, applied to those functions in $\mathfrak{Y}_0$ whose normal derivatives vanish on B, has a representation

$$-\nabla^2 = \int_0^\infty \mu\, dE_\mu,$$

where E_μ is a resolution of the identity. But then f_2 is given by the formula,

$$f_2 = \int_{0^-}^\infty [(\omega - i\lambda)^2 - \mu]^{-1}\, dE_\mu \{(2i\omega + \lambda) f_1 + g_1\}.$$

If λ is positive, the norm of the operator

$$\int_{0^-}^\infty [(\omega - i\lambda)^2 - \mu]^{-1}\, dE_\mu$$

is clearly bounded by $(2\omega\lambda)^{-1}$; and hence, the $L^{(2)}$ norms of f_2 and g_2, for small positive λ, are $\mathcal{O}(1/\lambda)$. This shows that the $L^{(2)}$ norms of v and v_t are bounded [3; p. 240].

Lemma 3. *If $u(x,t)$ is the solution of the homogeneous problem and $v(x,t) = u(x,t)\exp(-i\omega t)$, then*

$$w(x,T) = \int_0^T v(x,t)\,dt = \mathcal{O}(T^{\frac{1}{2}})$$

uniformly for x in V.

Proof. It is easy to verify that $v(x,t)$ satisfies the equation

$$(\nabla^2 + \omega^2)\,v(x,t) - 2i\omega v_t - v_{tt} = 0.$$

Thus,

$$(\nabla^2 + \omega^2)\,w(x,T) = [2i\omega v(x,t) + v_t(x,t)]_0^T = h(x,T).$$

By Lemma 1, both $w(x,T)$ and $h(x,T)$ have their supports in a sphere of radius $R+T$, and the $L^{(2)}$ norm of $h(x,T)$ is bounded. Because $w(x,T)$ has compact support, it satisfies *a priori* the Sommerfeld radiation condition. Moreover, there exists a Green's function $G(x,x')$ of the form [17]

$$G(x,x') = \frac{e^{i\omega|x-x'|}}{|x-x'|} + G^*(x,x')$$

for the exterior problem, where $G^*(x,x')$ has no singularities and is $\mathcal{O}(1/|x|)$ for large $|x|$; and $G(x,x')$ is such that

$$w(x,T) = \iiint_V \frac{e^{i\omega|x-x'|}}{|x-x'|}\cdot h(x',T)\,dx' + \iiint_V G^*(x,x')\,h(x',T)\,dx'.$$

If $|x| > R+T$, $w(x,T)$ vanishes; and hence we may assume that $|x| \leq R+T$. Let us compute the first integral over the intersection of V with a sphere with center at x and radius $2(R+T)$, and let us compute the second integral over the intersection of V and a sphere with center at the origin and radius $R+T$. Since both these spheres contain the support of $h(x,T)$, this replacement is permissible. But then, by using the Cauchy-Buniakovsky-Schwarz inequality, we find that both integrals are $\mathcal{O}(T^{\frac{1}{2}})$, which completes the proof.

We are indebted to Professor CALVIN WILCOX for a simplification in the proof of this lemma.

Proof of the Theorem. We begin by representing $\mathrm{col}(v, v_t)$, now determined by the inhomogeneous problem (I), by the variation of parameters formula

$$\begin{pmatrix} v \\ v_t \end{pmatrix} = e^{\mathfrak{m}t}\begin{pmatrix} f^* \\ g^* \end{pmatrix} + \int_0^t e^{\mathfrak{m}\xi}\begin{pmatrix} 0 \\ -\varrho \end{pmatrix} d\xi.$$

If we integrate both members of this equation with respect to T, we find that

$$\frac{1}{T}\int_0^T \begin{pmatrix} v \\ v_t \end{pmatrix} dt = \frac{1}{T}\int_0^T e^{\mathfrak{m}t}\begin{pmatrix} f^* \\ g^* \end{pmatrix} dt + \int_0^T \left(1 - \frac{\xi}{T}\right) e^{\mathfrak{m}\xi}\begin{pmatrix} 0 \\ -\varrho \end{pmatrix} d\xi.$$

The upper component in the first integral of the right-hand member is $\mathcal{O}(T^{-\frac{1}{2}})$ by Lemma 3. It therefore converges to zero uniformly as $T\to\infty$. Consequently, we lose no generality in assuming that $f^*=g^*=0$. Let $W(x, T)$ be the upper component of the left-hand member of the above equation, and let $w(x,\xi)$ be the upper component of $\exp(\mathfrak{m}\,\xi)\,\mathrm{col}(0,-\varrho)$. In particular,

$$(\nabla^2+\omega^2)\,w(x,\xi)-2i\omega\,w_\xi-w_{\xi\xi}=0,\quad w(x,0^+)=0,\quad w_\xi(x,0^+)=-\varrho.$$

Hence,

$$W(x,T)=\int_0^T\left(1-\frac{\xi}{T}\right)w(x,\xi)\,d\xi,$$

and

$$\begin{aligned}(\nabla^2+\omega^2)\,W(x,T)&=\int_0^T\left(1-\frac{\xi}{T}\right)[2i\omega\,w_\xi+w_{\xi\xi}]\,d\xi,\\&=\varrho(x)+\frac{2i\omega}{T}\int_0^T w(x,\xi)\,d\xi+\frac{1}{T}\,w(x,T).\end{aligned}$$

If $G(x,x')$ is the Green's function introduced in the proof of Lemma 3,

$$\begin{aligned}W(x,T)=&\iiint_V G(x,x')\,\varrho(x')\,dx'+\frac{2i\omega}{T}\iiint_V G(x,x')\int_0^T w(x',\xi)\,d\xi\,dx'+\\&+\frac{1}{T}\iiint_V G(x,x')\,w(x',T)\,dx'.\end{aligned}$$

As in Lemma 3, the integrals on the right-hand side of the above equation which involve T are seen to converge to zero as $T\to\infty$. Thus, the existence of

$$v^*(x)=\lim_{T\to\infty}W(x,T)=\iiint_V G(x,x')\,\varrho(x')\,dx'$$

has been shown. This proves (a) and (b).

A standard Abelian theorem [*1*; p. 193] and the observation that

$$\int_0^\infty v(x,t)\,e^{-st}\,dt=\int_{|x|-R}^\infty v(x,t)\,e^{-st}\,dt$$

imply (c) and (d). The proof of (e) rests on first observing, as we have, that the initial conditions can be ignored and then on obtaining the partial differential equation for $\hat{v}(x,s)$ from that satisfied by $v(x,t)$ by means of the usual methods of Laplace transform theory. This completes the proof.

Henceforward we shall consider only the reduced problem described in part (b) of the Theorem. Physically, this is not an unreasonable thing to do, since one can measure only the time average of a varying quantity, not its instantaneous values. It turns out that it is easiest to solve the reduced problem for $v^*(x)$ by determining the solution $\Phi(x,s)$ of (1) and then letting $s\to 0^+$. In order to

proceed further, we make the inevitable assumption that B is a level surface, $\xi=\xi_0$, in a coordinate system (ξ,η,φ) in which the operator $\nabla^2+(\omega-is)^2$ is separable. For simplicity, we assume further that B is symmetric in φ so as to reduce the number of variables which need be considered to two. Under these assumptions, then, equation (1) may be written in the form

$$-(L_\xi+L_\eta)\,\Phi=J(\xi,\eta)\,\varrho(\xi,\eta)$$

with

$$\partial\Phi(\xi_0,\eta)/\partial\xi=0$$

and where the operators L_ξ and L_η depend upon ξ and η alone, respectively.

The operator $L_\xi-\mu$ is not self-adjoint so that classical Sturm-Liouville theory cannot be used to find a representation for Φ. However, SIMS [*16*] and PHILLIPS [*14*] have provided a resolvent theory for operators of the same type as $L_\xi-\mu$ and $L_\eta-\nu$. Their work guarantees that there exist resolvents $\mathscr{R}_\mu$ and $\widetilde{R}_\nu$ for $L_\xi-\mu$ and $L_\eta-\nu$, respectively, with the property that for certain paths Γ in the μ-plane,

$$\Gamma: l+ics$$

$(-\infty<l<\infty$, c a positive constant depending on ω and the dimensions of $B)$,

$$\Phi(\xi,\eta,s)=\frac{1}{2\pi i}\int\limits_\Gamma \widetilde{R}_{-\mu}\mathscr{R}_\mu[J\varrho]\,d\mu.$$

A path Γ separates the poles of $\widetilde{R}_{-\mu}$ and $\mathscr{R}_\mu$. If ϱ reduces to a point source at (Ξ,τ) in the (ξ,η)-plane (a line source in cylinder problems), this integral simplifies and becomes

$$\Phi(\xi,\eta,\Xi,\tau,s)=\frac{1}{2\pi i}\int\limits_\Gamma \widetilde{G}(\eta,\tau,-\mu)\,G(\xi,\Xi,\mu)\,d\mu,$$

where $\widetilde{G}$ and G are the resolvent Green's function for the operators L_ξ and L_η, respectively.

One may attempt to evaluate this integral as a residue series by considering either the poles of $\widetilde{G}$ or those of G. The expansion involving the poles of the angular Green's function is a classical one, often called the Mie series expansion; but it converges so slowly when λ is small as to be almost useless then. The expansion involving the poles of the radial Green's function corresponds to the use of the Watson transform, familiar in right circular cylinder and sphere problems. However, except in such simple cases the proper formulation of the Watson transform had not been found prior to our approach. In cases which have been studied thus far, the radial residue series is rapidly convergent, when λ is small for points (ξ,η) in the geometric shadow of B.

It is in determining the terms of the radial residue series, particularly in locating the poles of G, that turning-point problems arise. To see this let us consider some specific examples.

A) *The elliptic cylinder.* Let B be an elliptic cylinder with semi-major axis a, semi-minor axis b, and eccentricity $e = \operatorname{sech} \xi_0$. In this case,

$$-L_\xi = \frac{d^2}{d\xi^2} + \gamma^2 \sinh^2 \xi$$

and

$$-L_\eta = \frac{d^2}{d\eta^2} + \gamma^2 \sin^2 \eta,$$

where $\gamma = ae(\omega - is)$. Then if $0 < \Im\mu < 2\omega b^2 s$, if y_1 is a solution of $(L_\xi - \mu) y = 0$ such that $y_1'(\xi_0) = 0$, and if y_2 is a solution in $L^{(2)}(\xi_0, \infty)$,

$$G(\xi, \Xi, \mu) = \frac{-1}{2i\gamma\, y_2'(\xi_0, \mu)} \begin{cases} y_1(\xi)\, y_2(\Xi), & \xi < \Xi \\ y_1(\Xi)\, y_2(\xi), & \Xi < \xi. \end{cases} \tag{2}$$

The poles of G are thus the zeros, in the μ-plane, of $y_2'(\xi_0, \mu)$.

Now y_2 is an $L^{(2)}$-solution of

$$\frac{d^2 y}{d\xi^2} + (\gamma^2 \sinh^2 \xi + \mu)\, y = 0 \qquad (\xi_0 \leqq \xi < \infty). \tag{3}$$

The conditions that λ be small and e be positive mean that equation (3) is to be considered for large $|\gamma|$. Note that as $e \to 1$, $\xi_0 \to 0$. If ξ_0 is bounded from zero, then the differential equation (3) has a simple turning point at ξ_0 when $\mu = -\gamma^2 \sinh^2 \xi_0$. One can show that for much smaller values of $|\mu|$, y_2 is not oscillatory. Therefore to find the zeros μ_n of $y_2'(\xi_0, \mu)$ one makes the substitution $\mu = -\gamma^2 \sinh^2 \xi_1$ and determines ξ_1 so that y_2' vanishes. This is done with the aid of LANGER's theory [*8*] for differential equations with simple turning points. The equation which determines μ_n is found to be [*5*]

$$\frac{d}{d\zeta} [\zeta^{\frac{1}{3}} H_{\frac{1}{3}}^{(2)}(\zeta)]_{\xi = \xi_0} + \mathcal{O}(\gamma^{-\frac{1}{3}}) = 0. \tag{4}$$

Here $\zeta = \gamma \int\limits_{\xi_1}^{\xi} (\sinh^2 t - \sinh^2 \xi_1)^{\frac{1}{2}} dt$ and $H_{\frac{1}{3}}^{(2)}$ is a Hankel function. It is possible to solve asymptotically the transcendental equation (4) for ξ_1 and thereby to obtain the functional dependence of μ_n on e and λ, provided e is bounded away from 1.

On the other hand, if e is so close to 1 that $\sinh^2 \xi_0 = \mathcal{O}(\gamma^{-1})$, then the zeros of $y_2'(\xi_0, \mu)$ may be located by means of MCKELVEY's second order turning point theory [*12*]. In this instance, one lets $\mu = -\gamma K$ and considers the equation (3) in the neighborhood of $\xi = 0$. The transcendental equation for μ_n which corresponds to (4) is considerably more complicated*.

The most important problem, however, is to derive the behavior of μ_n as $e \to 1$. To do this one needs to know the behavior of $y_2'(\xi, \mu)$ as a function of the

* An account of this work will be given in a forthcoming University of Michigan Research Institute Report.

vanishing 2Δ between the two confluent turning points of the differential equation

$$\frac{d^2 y}{d\xi^2} + \gamma^2(\sinh^2 \xi - \sinh^2 \Delta)\, y = 0. \tag{5}$$

This problem is presently being attacked via the theory in [6], which may be modified so as to apply to equation (5). Interest in the problem is pronounced because of the fact that as $e \to 1$, λ being fixed, slight changes in e produce relatively large changes in observed fields; that is to say, there is a resonance phenomenon. Moreover, none of the heuristic theories used by physicists and specialists in electromagnetic theory scattering problems are applicable in the resonance region.

Of course, after one has found the μ_n in these various situations, one must also derive the form of $\widetilde{G}(\eta, \tau, -\mu_n)$. This too requires turning point analyses of the sort described above.

B) *The prolate spheroid.* This case is more complicated than that of the elliptic cylinder as it is essentially three-dimensional. The complication is reflected in singularities in the operators L_ξ and L_η. Let B be a prolate spheroid with semi-major axis a, semi-minor axis b, and eccentricity $e = \xi_0^{-1}$. The operators L_ξ and L_η are then defined by the relations

$$-L_\xi\, y = [(\xi^2 - 1)\, y_\xi]_\xi + \gamma^2(\xi^2 - 1)\, y \qquad (\xi \geqq \xi_0)$$

and

$$-L_\eta\, y = [(1 - \eta^2)\, y_\eta]_\eta + \gamma^2(1 - \eta^2)\, y \qquad (-1 < \eta < 1).$$

As before, $\gamma = a e(\omega - i s)$ and $|\gamma|$ is large. If solutions y_1 and y_2 of $(L_\xi - \mu)\, y = 0$ are determined as in the elliptic cylinder case, $G(\xi, \Xi, \mu)$ is again defined by (2). Thus the problem reduces to one of locating the zeros μ_n of $y_2'(\xi_0, \mu)$, where y_2 is a solution of the equation

$$[(\xi^2 - 1)\, y_\xi]_\xi + [\gamma^2(\xi^2 - 1) + \mu]\, y = 0 \tag{6}$$

and y_2 lies in $L^{(2)}(\xi_0, \infty)$. Equation (6) has both a regular singularity and simple turning points. It ξ_0 is bounded away from 1, and is bounded, that is, if e is bounded away from both zero and 1, then in determining the early μ_n one may assume that $\mu - \mathcal{O}(\gamma^2)$ and use LANGER's theory [8] for a simple turning point. The evaluation of $\widetilde{G}$ requires knowledge of solutions of $(L_\eta - \mu)\, y = 0$ near $\eta = -1$. Since this is a singular point of the differential equation, a different asymptotic scheme of LANGER's [9] is used to evaluate $\widetilde{G}$. This analysis was carried out in [4]. If $(\xi - 1) = \mathcal{O}(\gamma^{-1})$, then the above mentioned study of $(L_\xi - \mu)\, y = 0$ no longer applies. The equation $(L_\xi - \mu)\, y = 0$ must be considered in the neighborhood of its singularity at $\xi = 1$, and the turning point no longer plays a rôle.

To obtain the asymptotic behavior of solutions of $(L_\xi - \mu)\, y = 0$ near $\xi = 1$ one uses MCKELVEY's theory [13]. This analysis is not nearly as complicated, which is natural since it corresponds to the case of B being a needle; and a needle disturbs the incident radiation very little.

In the region between $(\xi_0^2-1)=\mathcal{O}(\gamma^{-1})$ and $(\xi_0^2-1)=\mathcal{O}(1)$, the resonance region, the effect upon $y_2(\xi,\mu)$ of both the turning point and singularity of $(L_\xi-\mu)\,y=0$ must be considered. In particular, the problem is to determine the functional behavior of y_2 as these points coalesce. Although the situation in the limits is simple, the simple pole and zero cancel so that there is neither a turning point nor a singularity in the coefficient of γ^2 to cause difficulty, no existing asymptotic theory appears to be adaptable to the solution of this problem.

The above examples show that the field of diffraction theory is as rich a one for the application of the theory of asymptotic solutions of ordinary differential equations with turning points as has yet been found. It demands all aspects of the theory and points out the need for new extensions. In case B is not a level surface of a coordinate system in which the operator $\nabla^2+(\omega-is)^2$ is separable, neither the generalized Sturm-Liouville theory nor the turning-point theory we have discussed here applies. The development of satisfactory theories to replace them is perhaps the most challenging and difficult problem in this field.

Part of this work was done under Contract AF-19(604)-4993, Air Force Cambridge Research Center, ARDC.

References

[1] Doetsch, G.: Theorie und Anwendung der Laplace-Transformation. Berlin: Springer 1937.

[2] Fock, V. A.: Diffraction of radio waves around the earth's surface. J. Physics **9**, 255–266 (1955).

[3] Hille, E.: Functional analysis and semi-groups. Amer. Math. Soc. Colloquium Publications XXXI, Amer. Math. Soc., Providence, 1948.

[4] Kazarinoff, N. D., & R. K. Ritt: On the theory of scalar diffraction and its application to the prolate spheroid. Annals of Physics **6**, 277–299 (1959).

[5] Kazarinoff, N. D., & R. K. Ritt: Scalar diffraction by an elliptic cylinder. University of Michigan Research Institute Rept. 2871-2-T (1959).

[6] Kazarinoff, N. D.: Asymptotic theory of second order differential equations with two simple turning points. Archive for Rational Mech. and Anal. **2**, 129–150 (1958).

[7] Ladyženskaya, O. A., & M. I. Visik: Boundary value problems for partial differential equations and certain classes of operator equations. Uspehi Mat. Nauk (N.S.) **11**, 41–97 (1956).

[8] Langer, R. E.: The asymptotic solutions of ordinary linear differential equations of the second order with special reference to a turning point. Trans. Amer. Math. Soc. **67**, 461–490 (1949).

[9] Langer, R. E.: On the asymptotic solutions of ordinary differential equations, with reference to the Stokes phenomenon about a singular point. Trans. Amer. Math. Soc. **37**, 397–416 (1935).

[10] Levy, B. R., & J. B. Keller: Diffraction by a smooth object. Comm. on Pure and Appl. Math. **12**, 159–209 (1959).

[11] Lewis, R. M.: Discontinuous initial value problems for symmetric hyperbolic linear differential equations. J. of Math. and Mech. **7**, 571–592 (1958).

[12] McKelvey, R.: The solutions of second order linear ordinary differential equations about a turning point of order two. Trans. Amer. Math. Soc. **79**, 103–123 (1959).

[13] McKelvey, R.: Solution about a singular point of a linear differential equation involving a large parameter. Trans. Amer. Math. Soc. **91**, 410–424 (1959).

[14] Phillips, R. S.: Linear ordinary differential operators of the second order. N.Y.U. Research Report EM-42 (1952).

[*15*] RITT, R. K.: Studies in Radar Cross Sections XXX. University of Michigan Research Institute Rept. 2591-4-T (1958).
[*16*] SIMS, A. R.: Secondary conditions for linear differential operators of the second order. J. of Math. and Mech. **6**, 247—285 (1957).
[*17*] WEYL, H.: Die natürlichen Randwertaufgaben im Außenraum für Strahlungsfelder beliebiger Dimension und beliebigen Ranges. Math. Z. **56**, 105—119 (1952).
[*18*] WILCOX, C. H.: Spherical means and radiation conditions. Arch. Rational Mech. Anal. **3**, 133—148 (1959).

The University of Michigan
Ann Arbor, Michigan
and
Mathematics Research Center
United States Army
Madison, Wisconsin

(Received February 3, 1960)

EDITORIAL BOARD

CONTENTS

Druck der Universitätsdruckerei H. Stürtz AG., Würzburg
Printed in Germany